

## THESIS / THÈSE

### MASTER EN INGÉNIEUR DE GESTION À FINALITÉ SPÉCIALISÉE EN DATA SCIENCE

#### Vers une approche par pattern de la conception de Data Warehouses?

de Munck, Noémie

*Award date:*  
2021

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Vers une approche par pattern de la conception de Data Warehouses ?

**Noémie DE MUNCK**

**Directeur: Prof. I. LINDEN**

Mémoire présenté  
en vue de l'obtention du titre de  
Master 120 en ingénieur de gestion, à finalité spécialisée  
en data science

**ANNEE ACADEMIQUE 2020-2021**

UNIVERSITÉ DE NAMUR

MÉMOIRE DANS LE CADRE DU MASTER EN INGÉNIEUR DE  
GESTION À FINALITÉ SPÉCIALISÉE

FACULTÉ DES SCIENCES ÉCONOMIQUES, SOCIALES ET DE GESTION

---

**Vers une approche par pattern de la conception de Data  
Warehouses?**

---

**RÉSUMÉ**

Les Data Warehouses sont un des éléments clés d'une solution de Business Intelligence. Pour concevoir leur structure, il faut suivre des méthodes formelles. Cependant, l'application de ces méthodes peut présenter des difficultés. Dans le domaine de l'ingénierie logicielle, le concept de design patterns a fait ses preuves en matière de résolution de problèmes de conception. Ce concept propose des solutions pour résoudre des problèmes courants de développement. On peut considérer les conceptions existantes de Data Warehouses comme des solutions éprouvées. L'objectif de ce mémoire sera donc d'explorer la possibilité de design patterns au cas de la conception de Data Warehouses et d'en discuter les critères d'applicabilité.

Auteur : Noémie DE MUNCK

Promoteur : Isabelle LINDEN

Juin 2021

## **Avant-propos**

Je tiens à remercier toutes les personnes qui m'ont accordé du temps et qui ont joué un rôle important dans la réalisation de ce travail.

En premier lieu, je remercie ma promotrice de mémoire, Isabelle Linden, qui m'a permis de choisir un sujet intéressant et d'actualité. En effet, la problématique de la Business Intelligence et notamment la conception de ses Data Warehouses sont au centre des préoccupations de nombreuses entreprises ou organisations. De plus, je tiens à la remercier d'avoir été présente pour encadrer mon travail, répondre à mes questions et m'aiguiller dans la réalisation de ce mémoire.

Ensuite, je remercie ma sœur, Delphine de Munck de m'avoir aidé pour la mise en page et rédaction de mon mémoire. Je tiens également à remercier mon papa, Joël de Munck, pour avoir toujours été présent pour discuter de mon mémoire et m'aider à trouver des pistes de solutions. Je souhaite aussi remercier ma maman, Sophie Fayt, pour ses nombreuses relectures attentives de mon mémoire ainsi que pour ses nombreuses pistes d'améliorations. Je les remercie également tous les trois d'avoir toujours été là pour moi durant mes études et d'avoir toujours cru en moi.

De plus, je tiens à remercier mes amies proches et mon compagnon, Sarah Lebrun, Jessica Degueldre et Maxime Gournis, qui, ayant tous les trois fait les mêmes études à des années différentes, m'ont toujours guidé, motivé, épaulé, écouté et soutenu chacun à leur manière.



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Revue de la littérature</b>	<b>5</b>
1.1 Concept de design patterns . . . . .	5
1.2 Software design patterns . . . . .	6
1.3 Database design patterns . . . . .	8
<b>2 Méthodologie</b>	<b>9</b>
2.1 Collecte de schémas de conception . . . . .	9
2.2 Analyse des schémas collectés . . . . .	10
2.3 Identification de patterns . . . . .	10
2.4 Méthodologie proposée . . . . .	10
<b>3 Analyse</b>	<b>11</b>
3.1 Collecte de schémas de conception . . . . .	11
3.2 Analyse des schémas collectés . . . . .	13
3.3 Application du concept de patterns de design . . . . .	20
3.3.1 Adaptation du concept . . . . .	21
3.3.2 Lien avec l'identification des exigences . . . . .	22
3.3.3 Identification des design patterns de domaine . . . . .	23
3.4 Design pattern de domaine . . . . .	23
3.4.1 Domaine du commerce . . . . .	24
3.4.2 Domaine de la gestion . . . . .	30

3.4.3	Domaine des télécommunications . . . . .	34
3.4.4	Domaine société . . . . .	38
3.4.5	Domaine de l'éducation . . . . .	44
3.4.6	Domaine médical . . . . .	48
3.4.7	Domaine de la culture . . . . .	53
3.4.8	Domaine de la géologie . . . . .	58
<b>4</b>	<b>Résultats</b>	<b>62</b>
4.1	Application . . . . .	62
4.2	Limites . . . . .	65
4.3	Discussion . . . . .	66
<b>5</b>	<b>Conclusion</b>	<b>68</b>
	<b>Annexe</b>	<b>69</b>
	<b>Références</b>	<b>158</b>

# Introduction

## Domaine d'étude

Il est de plus en plus reconnu que les données sont l'un des éléments les plus importants de toute organisation [1]. Conscientes de cette importance, les entreprises ont commencé, depuis la fin des années 1970, à stocker leurs données dans des bases de données relationnelles. Depuis lors, la quantité d'informations collectées n'a cessé d'augmenter. De plus, elles sont maintenant dispersées entre différents systèmes, sous formes diverses et variées et même parfois à des endroits géographiques différents. Les entreprises, faisant face à une dynamique de marché et une compétition accrue, ont plus que jamais besoin de prendre les bonnes décisions sur les bonnes informations au bon moment. L'analyse d'une telle quantité de données étant compliquée sans aide, les entreprises ont donc besoin d'outils d'analyse pour les soutenir dans leur prise de décision [2]. C'est ici qu'intervient la Business Intelligence.

La Business Intelligence comprend un ensemble de méthodes et technologies qui permettent de transformer des données brutes en informations significatives pour la prise de décision. Son but est de fournir une assistance aux responsables d'organisations pour analyser leurs données et informations stratégiques. Pour ce faire, la Business Intelligence collecte les données de la société, les intègre et les transforme de manière à ce qu'elles puissent être utilisées pour analyser le comportement de l'entreprise en question. Lorsque les données ont été extraites de leurs différentes sources, transformées, intégrées et nettoyées, elles sont alors stockées dans une nouvelle base de données. Cette nouvelle base de données s'appelle un Data Warehouse [2].

Les Data Warehouses sont donc l'élément clé d'une solution de Business Intelligence. Ils permettent de rassembler les données de sources d'information différentes et de types différents dans un seul entrepôt adapté à l'analyse de celles-ci. Ils ont été développés et déployés pour fournir une infrastructure qui permet aux utilisateurs d'obtenir des

réponses efficaces et précises à des requêtes complexes. Beaucoup de systèmes et d'outils peuvent être utilisés pour avoir accès, analyser et exploiter les données contenues dans ceux-ci [2].

Les Data Warehouses peuvent être vus comme un type de base de données. Cependant, par rapport aux bases de données opérationnelles, les Data Warehouses sont assez différents surtout du point de vue du développement. En effet, il y a lieu de représenter les Data Warehouses selon une modélisation multidimensionnelle et non relationnelle [2]. Les Data Warehouses ont dès lors des ouvrages de référence qui leur sont propres.

Parmi ces ouvrages de référence, on distingue trois livres qui font autorité dans la communauté scientifique. Le premier est un livre de Ralph Kimball et Margy Ross intitulé « The Data Warehouse Toolkit – Third Edition » [1]. Le deuxième est un livre d'Alejandro Vaisman et Esteban Zimanyi intitulé « Data Warehouse Systems : Design and Implementation » [2]. Le troisième est un ouvrage de Matteo Golfarelli et Stefano Rizzi intitulé « Data Warehouse Design : Modern Principles and Methodologies » [3].

Ces ouvrages fort complets et précis reprennent les différents aspects et nuances que peuvent comporter les Data Warehouses et abordent également leur domaine d'application et des aspects plus avancés. Ils sont des références clés pour la conception de Data Warehouses.

## **Sujet de recherche**

Bien que ces livres soient d'une qualité incontestable, baser le développement de Data Warehouses uniquement sur ceux-ci peut s'avérer compliqué. L'appropriation des méthodes présentes dans ces livres demande un travail significatif. En outre, si l'on suit à la lettre les méthodologies formelles qu'ils décrivent, la réalisation risque d'être assez longue. De plus, dans de nombreux domaines, la création de Data Warehouse ne nécessite pas de maîtriser toute la complexité théorique de la tâche.

On peut constater que de nombreux articles scientifiques [4–13] ont également iden-

tifié ce problème de conception. Tous ces articles expriment des frustrations quant aux méthodologies actuelles de conception des Data Warehouses. La plupart de ces articles, neuf des dix analysés, ont mentionné que la réponse la plus adéquate à ce problème serait de créer une nouvelle méthode de conception. En effet, ces neuf articles ont chacun à leur manière créé soit un cadre méthodologique de conception, soit une approche conceptuelle, soit une méthode de conception, soit un modèle conceptuel, soit une liste de points d'attention pour la construction. En résumé, on peut donc constater que ces articles débattent sur les lignes directrices à suivre lors de la conception d'un Data Warehouse. De plus, en regardant plus attentivement ce que les auteurs indiquent dans leurs différentes méthodes, on peut constater que beaucoup d'éléments sont assez semblables. De surcroît, on remarque que ces éléments sont également très similaires à ceux des méthodes des trois livres de références majeurs [1–3].

En outre, les différences qui peuvent exister entre ces multiples méthodes sont normales. De fait, suivant les préférences de chacun et les cas d'applications qui varient, il est logique que des différences existent. La conception de Data Warehouses étant relativement dépendante de son créateur, il est normal que le même cas d'étude donnera différents schémas de Data Warehouses. L'élément principal réside dans la présence des grandes étapes de conception et de solutions semblables.

Dans le domaine de la Business Intelligence, on peut donc établir que la conception des Data Warehouses est une problématique bien établie. De plus, on constate que les différentes solutions existantes laissent place à des pistes d'exploration. En effet, il semble que le sujet n'a pas été complètement épuisé. Pour tenter d'apporter une nouvelle solution à cette problématique, une idée pourrait être de s'inspirer de la résolution de problèmes de manière générale. En effet, certaines méthodologies ont fait leurs preuves dans d'autres domaines. C'est le cas des patterns de design. Pourquoi ce concept ne pourrait-il pas être adapté au cas de la conception des Data Warehouses ? Plus concrètement, la question de recherche de ce mémoire est donc : « Une méthode de design basée sur des patterns pourrait-elle être proposée pour la conception de Data Warehouses ? ».

Pour tenter de répondre à cette question, ce mémoire est structuré comme suit. Dans un premier temps, il est judicieux de s'intéresser de plus près au concept de design orienté pattern. Ce travail commence donc par une revue de la littérature scientifique sur le sujet des patterns de design et plus particulièrement leur application au domaine de l'ingénierie logicielle et à celui des bases de données où ils font l'objet d'une large littérature. Par après, la méthodologie proposée dans ce travail pour explorer l'adaptation du concept au cas de la conception des Data Warehouses est expliquée. La partie suivante est relative à l'analyse mettant en pratique la méthodologie proposée précédemment. Cette analyse consiste à collecter et analyser des conceptions existantes de Data Warehouses dans le but d'explorer les possibilités d'identification et d'application du concept. Par la suite, les résultats de cette analyse sont discutés en terme d'application, de limites et d'améliorations. Finalement, une conclusion générale est réalisée en reprenant les points essentiels abordés dans ce travail.

# 1 Revue de la littérature

Les patterns de design sont largement exploités dans l'ingénierie informatique notamment pour le design de logiciels et de bases de données. Avant d'envisager une telle démarche pour le design de Data Warehouses, ce chapitre fait le point sur le concept de pattern de design et la mise en oeuvre dans les domaines des logiciels et des bases de données.

## 1.1 Concept de design patterns

La théorie des patterns a été proposée et perfectionnée par Christopher Alexander [14]. Un des points forts de ce concept est d'être applicable à tous types de processus de conception. Par exemple et étonnement, c'est dans le domaine de l'histoire de l'art et de l'anthropologie que ce concept a tout d'abord été utilisé [15]. Notre intérêt se porte sur l'usage de ce concept dans le domaine des bases de données et dans celui de l'ingénierie logicielle. Ce concept étant facilement applicable à différents domaines, cela renforce encore plus l'idée de l'appliquer au cas de la conception de Data Warehouses.

L'utilisation de patterns est une pratique permettant de résoudre des problèmes courants. Pour ceux-ci, des modèles peuvent être utilisés pour faciliter la réutilisation de solutions déjà bien établies. Ces modèles proposent donc des solutions pour résoudre des problèmes généraux de conception dans des contextes particuliers. Pour ce faire, ils capturent des solutions qui ont déjà été développées et éprouvées au fil du temps. Plus particulièrement, ils décrivent les structures récurrentes des composants permettant de résoudre des situations données [16].

Ce concept de design patterns a prouvé être un moyen utile pour capturer des solutions de conception testées dans le temps et de faciliter leur réutilisation. En d'autres termes, ces patterns sont des schémas de solutions éprouvées [14]. Il s'agit d'une description de la manière d'aborder ces problèmes et de concevoir une solution [17]. Le but

des patterns est de représenter explicitement les connaissances de conception qui sont connues implicitement par les praticiens qualifiés [18].

## 1.2 Software design patterns

Dans le domaine de l'ingénierie de logiciels, l'utilisation de design patterns permet de disposer de solutions récurrentes pour des problèmes de logiciel standards. L'industrie moderne de logiciels a adopté cette méthodologie de design patterns pour permettre la réutilisation des bonnes pratiques. Cette manière de procéder permet d'améliorer la qualité des systèmes logiciels [19]. C'est sans doute dans ce domaine que le concept de design patterns est le plus efficace [18].

En général dans la littérature récente de ce domaine, les descriptions mettent plus l'accent sur les nouvelles caractéristiques et fonctionnalités que sur les modèles de conception les plus récurrents [18]. Pourtant, certaines problématiques de conception sont semblables et peuvent être identifiées par les développeurs. Conscients qu'il n'est pas utile de réinventer la roue à chaque fois, ceux-ci ont donc jugé opportun de conceptualiser ces motifs de conception pour permettre de les réutiliser [20]. Les design patterns ont été créés pour disposer d'un seul outil approprié pour les problématiques récurrentes d'un même domaine [18].

L'utilisation de design patterns dans le domaine de l'ingénierie logicielle orientée objet consiste en la description d'objets et de classes communicants qui peuvent être personnalisés pour résoudre des problèmes de conception dans un contexte particulier. Ces conceptualisations se traduisent donc dans des idées abstraites concernant la résolution d'un problème récurrent et non par du code directement [20]. Cela signifie qu'un design pattern représente une idée dans un paradigme spécifique mais pas une implémentation particulière. En effet, les patterns documentent les solutions de conception de logiciels orientés objet indépendamment des différents langages de programmation [21]. Typiquement, un pattern est composé d'un nom, d'un emploi, d'une description concernant



comment et quand l'utiliser, d'un diagramme structurel, d'exemples d'utilisation et d'une discussion des différentes interactions avec les autres patterns [18].

Il existe différents types de design patterns dont trois principaux. Tout d'abord, il y a les patterns relatifs à la structure. Ceux-ci sont des modèles qui concernent l'organisation des différentes classes et objets. Le deuxième type de design patterns est quant à lui relatif au comportement du logiciel. Ce type concerne l'identification des modèles de communication communs entre les objets. Le dernier type de design patterns est celui relatif à la création [14]. Les patterns de ce type concernent soit l'instanciation de classes soit la création d'objets [21].

Un cas d'usage concret est le design pattern adaptateur. Celui-ci est utilisé lorsque deux interfaces sont incompatibles mais que l'on souhaite tout de même établir une relation entre elles. Le pattern adaptateur convertit alors l'interface d'une classe en une autre interface ou classe selon ce que le client attend. L'adaptateur permet aux classes incompatibles de tout de même collaborer [21].

Un autre cas d'usage très courant est le client-serveur. Comme son nom l'indique, il se compose d'un serveur et de clients multiples. Le composant serveur fournit des services à plusieurs composants clients, les clients demandent des services au serveur et le serveur fournit les services demandés aux clients concernés. Ce pattern est utilisé lorsqu'un serveur et ses clients se connectent par internet [22]. Il est donc beaucoup utilisé dans les applications en ligne telles que les e-mails ou les documents partagés [23].

Le fait d'intégrer des design patterns au processus de développement présente des avantages. L'utilisation de patterns permet de réutiliser les architectures de conception de logiciels et d'améliorer la communication entre et à travers les équipes de développement. De plus, cela facilite sensiblement la formation de nouveaux programmeurs et aide à dépasser les différentes manières de penser imposées par les différents langages de programmation [18]. De surcroît, les design patterns permettent une programmation plus rapide et plus sûre [14]. En conséquence, l'utilisation de design patterns permet d'augmenter la flexibilité, l'extensibilité et la maintenabilité du système [24].

### 1.3 Database design patterns

Le concept des design patterns est également utilisé dans le domaine des bases de données. Cette utilisation est principalement due à la popularité grandissante de ce concept dans différentes disciplines et surtout dans le domaine de l'ingénierie logicielle. En comparaison avec ce dernier, les recherches, applications et utilisations de ce concept sont moins nombreuses et plus récentes. Nul ne fait mention de fondateur ou d'initiateur du concept à ce domaine. Il en va donc de même pour l'année précise de la première adaptation de ce concept aux bases de données.

Les design patterns commencent donc à être utilisés pour la conception de bases de données car il semble plus opportun de regarder comment les développeurs conçoivent leurs bases de données plutôt que de se demander comment ils devraient le faire. Pour y parvenir, la première étape est de rassembler le plus de conceptions de bases de données possible [25].

Du fait de la jeunesse du domaine, contrairement à l'ingénierie logicielle, tous les auteurs ne s'accordent pas sur la définition des design patterns de référence pour les bases de données. On constate que certains identifient trois patterns qui sont le pivotement, la matérialisation et la généralisation/spécialisation. D'autres en identifient deux principaux qui sont l'internationalisation et le registre des transactions accessibles. On remarque également que d'autres encore en identifient deux qui sont les rôles multiples et les hiérarchies. En résumé, on constate que l'application du concept de design patterns aux cas des bases de données nécessite encore certains ajustements [25].

Même si l'utilisation de design patterns dans le cadre de la conception de base de données n'est pas encore tout à fait au point, plusieurs auteurs soulignent que l'utilisation de ce concept présente beaucoup d'avantages [24–26]. Même pour des novices, travailler avec des design patterns permettrait une conception plus aisée et plus rapide. C'est pourquoi, certains pensent intégrer directement ce concept dans les cours de bases de données [24].

## 2 Méthodologie

Comme on a pu le constater, l'utilisation de design patterns amène beaucoup d'avantages. L'application de ce concept à l'ingénierie de logiciels n'est plus à démontrer. En ce qui concerne l'application aux bases de données, même si tous ne sont pas encore d'accord sur l'application à proprement parler, tous s'accordent sur le fait que cela présente beaucoup d'avantages. Il semble donc judicieux et possible d'appliquer ce concept au cas de la conception de Data Warehouses.

Pour tenter d'adapter ce concept de design patterns au cas de la conception de Data Warehouses, une méthodologie doit être mise en place. Celle-ci consiste en trois étapes. La première est la collecte de schémas de conceptions existantes. La deuxième étape est l'analyse de ces schémas de solutions existantes de conception. La dernière étape consiste en l'identification et la représentation des design patterns sur base des éléments identifiés par l'analyse.

### 2.1 Collecte de schémas de conception

Le concept de design patterns repose sur des solutions de design existantes et éprouvées. Celles-ci servent de base à l'identification de patterns de design. Il est donc primordial de disposer de telles solutions.

De nombreuses entreprises ayant adopté la Business Intelligence, des milliers de Data Warehouses ont déjà été construits et ce sur des sujets divers et variés [2]. Ces Data Warehouses déjà construits et éprouvés sont le point de départ en vue d'explorer la possibilité d'identification de patterns de design. En effet, ceux-ci seront les solutions éprouvées servant de fondation à l'application du concept. Elles se doivent donc d'être de bonne qualité. La collecte de celles-ci devra alors assurer la variété des sources ainsi que leur véracité.

## **2.2 Analyse des schémas collectés**

La deuxième étape de la méthodologie consiste à analyser les cas de conception récoltés à l'étape précédente. Cette analyse a comme but d'identifier les éléments étant déterminants pour la conception.

Cette étape est fondamentale pour l'exploration d'application du concept de design patterns. En effet, l'application dépendra des éléments identifiés. Ces éléments servent de pistes à l'identification de problématiques de conception bien précises où il pourrait être opportun d'analyser s'il existe des patterns dans la conception.

## **2.3 Identification de patterns**

Sur base des éléments mis en lumière à l'étape précédente, la recherche de patterns de design peut se faire. Chacun des éléments identifiés précédemment représente un candidat potentiel à la définition d'un pattern de design.

Dans cette étape, si des patterns de design sont identifiés, il sera également question de déterminer la manière la plus optimale de les représenter. Celle-ci doit permettre une compréhension aisée pour les utilisateurs. De plus, elle doit également permettre de faciliter l'utilisation et l'adaptation des patterns de design.

## **2.4 Méthodologie proposée**

Pour aboutir à la définition de patterns de référence, la première étape consiste à collecter des schémas de Data Warehouses. La deuxième étape de la méthodologie proposée est d'analyser ces schémas collectés dans l'optique d'identifier les éléments déterminants à leur conception. Sur base des éléments identifiés, la troisième étape consiste à explorer la possibilité des design patterns. Dans le cas où cette possibilité existe, il est alors également question de les identifier et de les représenter. Dans le chapitre suivant, une analyse correspondant à cette méthodologie est proposée.

## 3 Analyse

Pour tenter d'obtenir les design patterns souhaités, une analyse est nécessaire. Celle-ci s'organise en trois étapes. La première consiste à collecter des schémas existants de Data Warehouses. Ceux-ci peuvent être considérés comme des solutions éprouvées sur lesquelles on peut se baser pour identifier les design patterns. La deuxième étape est l'analyse de ces schémas. L'objectif de celle-ci est d'identifier quels éléments sont déterminants pour la conception de Data Warehouses. Enfin sur base de ces éléments, la troisième étape explore l'application du concept de design patterns. Si une telle application est possible, cette troisième étape réalise également l'identification et la représentation des design patterns.

### 3.1 Collecte de schémas de conception

Les schémas de conception étant la base sur laquelle repose le concept de pattern de design, ce recueil se doit d'être réalisé à travers des sources variées. En effet, cela permet de garantir plus de qualité. Trois différents types de sources seront donc utilisés pour rechercher les schémas de conception nécessaires à cette analyse.

Le premier type de sources est constitué des divers exemples et cas d'études présents dans les livres de références sur les Data Warehouses. Pour rappel, trois sortent du lot en terme de qualité [1–3]. Ayant été créés par des experts sur le sujet, les schémas de ces livres sont de très bonne qualité.

De ce premier type de sources, 35 schémas de Data Warehouses ont été recueillis. La table 11, disponible en annexe, détaille quels schémas ont été extraits de quels livres.<sup>1</sup>

Le deuxième type de sources utilisées pour recueillir des schémas de Data Warehouses est constitué des articles scientifiques. Pour continuer dans la lignée de la bonne qua-

---

1. Le numéro d'identification de chaque schéma ainsi que le numéro de page à laquelle on peut le retrouver en annexe sont indiqués dans la table. Cette table contient le nom des auteurs et le titre du livre d'où a été extrait chaque schéma. Il y est également indiqué le numéro de référence vers la bibliographie.

lité de schémas recueillis avec le premier type de sources, des critères ont été fixés pour déterminer quels articles peuvent être pris en compte pour réaliser cet inventaire de schémas de Data Warehouses. Ces critères concernent les auteurs des articles et sont au nombre de deux. Le premier critère concerne le nombre de publications des auteurs. En effet, il semble que cinq articles publiés soient un minimum. Le second critère concerne le fait que les auteurs de l'article doivent être reconnus dans le monde scientifique. Cela signifie qu'au minimum un des auteurs de chaque article doit soit être professeur d'université, soit détenir un doctorat.

Ce deuxième type de sources est celui qui permet de collecter le plus de schémas. En effet, après avoir appliqué les deux critères de sélection, 50 schémas ont été recueillis. Les informations concernant les différents schémas récoltés de ce type de sources sont disponibles dans la table 12, disponible en annexe.<sup>2</sup>

Le troisième et dernier type de sources d'où des schémas de Data Warehouses ont été extraits est constitué de sites internet. Il est évident que cette dernière source est a priori de qualité plus variable que les deux précédentes. Cependant, pour être le plus complet possible dans l'analyse, il semble important de ne pas s'arrêter uniquement aux pratiques du monde scientifique mais d'inclure également des schémas provenant de personnes pas forcément reconnues comme expertes en la matière. Malgré le fait que cette source soit d'une moins bonne qualité que les précédentes, certains critères peuvent être établis pour ne prendre uniquement que les meilleurs sites et de ce fait être plus constant dans la qualité des schémas recueillis. Trois critères ont été établis pour déterminer si un site internet pouvait être pris en compte pour cette analyse. Le premier critère est le fait que le nom de l'auteur soit disponible sur le site. En effet, les sites où le nom de l'auteur apparaît ont une meilleure probabilité d'être de qualité. Le deuxième critère concerne les explications fournies avec le schéma de Data Warehouses. De fait, un schéma accompagné d'explications logiques peut être considéré comme étant de meilleure qualité. Le

---

2. Cette table contient le numéro d'identification des schémas, le numéro de la page en annexe de laquelle ils se trouvent ainsi que le nom, les auteurs et la référence de l'article.

dernier critère concerne la présence d'une date de publication. En effet, la présence de la date de publication du site témoigne d'une fiabilité plus grande.

15 schémas de Data Warehouses ont été recueillis par cette troisième démarche. La table 13, disponible en annexe, détaille quels schémas ont été extraits de quel site.<sup>3</sup>

En résumé, de ces trois différents types de sources, 100 schémas ont été collectés. La répartition est de 35 en provenance des livres de références, 50 venant d'articles scientifiques et 15 issus de sites internet. Trois tableaux, disponibles en annexe, reprennent l'inventaire des schémas de Data Warehouses recueillis selon le type de sources. Tous les schémas sont disponibles en annexe.

### 3.2 Analyse des schémas collectés

Après la collecte des schémas de conception, la deuxième étape de notre étude est d'analyser ces différents schémas. Cette analyse porte sur divers éléments caractérisant les schémas. Celle-ci est synthétisée par une table et trois graphiques. Cette analyse permet de grouper les schémas selon huit domaines métier.

L'analyse de chaque schéma porte sur sa structure ainsi que sur les descriptions et autres explications disponibles dans le livre, l'article ou le site d'où il a été extrait. Cette analyse permet de mieux comprendre chaque schéma et d'extraire des informations importantes sur ceux-ci.

Pour tenter d'identifier au mieux des tendances dans la conception de Data Warehouses, il est opportun d'avoir une vue plus globale sur les schémas. Pour ce faire, l'analyse des schémas est synthétisée dans la table 1<sup>4</sup> qui permet de regrouper les informa-

---

3. Le numéro d'identification du schéma et le numéro de la page à laquelle on peut le retrouver sont y sont indiqués. On retrouve le nom des auteurs et le titre du site internet dont ils ont été extraits. Il y est également indiqué le numéro de référence vers la bibliographie.

4. Dans la table 1, plusieurs colonnes concernent la structure des schémas telles que le nombre de dimensions, le nombre de hiérarchies ou la forme. Il y a également une colonne avec le numéro du schéma et une avec le sujet central d'analyse. Pour les schémas ayant plusieurs tables de fait, ils sont décomposés selon les différents sujets d'analyse et repris plusieurs fois dans la table. Les schémas sont groupés par domaine métier et ordonnés par ordre décroissant du nombre de schémas de chaque domaine.

tions pertinentes en un seul endroit et de comparer plus facilement les schémas entre eux. De plus, les résultats de l'analyse sont également synthétisés dans les figures 1, 2 et 3<sup>5</sup> qui permettent de distinguer des tendances.

La classification des schémas suivant les domaines métier a été réalisée en trois étapes. La première est d'identifier le sujet central d'analyse des schémas. La deuxième étape consiste à regrouper les sujets portant sur des thèmes semblables. La dernière étape est l'identification du domaine métier représentant le mieux les sujets regroupés. Il est évident qu'il existe bien d'autres domaines métier où la Business Intelligence est, ou pourrait être, utilisée. Cependant, cette analyse se base uniquement sur les schémas collectés.

Huit domaines métier ont été identifiés. Le premier concerne le commerce et est composé des sujets relatifs aux ventes<sup>6</sup>, inventaires ou achats. Le second concerne la gestion qui peut être de budget ou encore d'activités. Le troisième est le domaine des télécommunications qui regroupe ce qui touche aux ordinateurs, à internet ainsi qu'à la téléphonie. On retrouve également un domaine relatif à tout ce qui concerne la vie en société. Le cinquième domaine regroupe les sujets relatifs à l'éducation tels que les inscriptions ou les évaluations des étudiants. Il y a aussi un domaine en rapport aux sujets médicaux. Le domaine de la culture se compose des sujets relatifs à la littérature, la presse ou le cinéma. Le dernier domaine est relatif à la géologie. Presque tous les schémas sont relatifs à ces huit domaines métier. Cependant, deux schémas ne portent pas sur des thèmes semblables. Ils ont alors été regroupés dans un domaine "autre".

---

5. La figure 1 représente le nombre de schémas par domaine métier. La figure 2 représente le nombre de schémas par domaine selon leur forme. La figure 3 représente le nombre de schémas par domaine et par nombre de dimensions.

6. La sur-représentation de la thématique des ventes est due au fait que le cas d'étude classique pour les Data Warehouses est celui qui analyse les ventes d'une société.



TABLE 1: Table récapitulative des caractéristiques des schémas

Domaine	Nombre	Numéro	Sujet	Dimension	Hiérarchie	Forme
Commerce	45	1	Vente	17	6	Flocon
		2	Vente	17	5	Flocon
		4	Vente	4	0	Etoile
		5	Vente	7	2	Flocon
		8	Vente	6	0	Etoile
		9	Inventaire	3	0	Etoile
		10	Inventaire	8	0	Etoile
		11	Achat	5	0	Etoile
		12	Achat	11	0	Etoile
		13	Commande	6	0	Etoile
		14	Export	8	0	Etoile
		26	Vente	20	10	Flocon
		27	Export	26	12	Flocon
		30	Vente	6	3	Flocon
		32	Inventaire	11	4	Flocon
		37	Client	7	0	Etoile
		44	Commande	6	2	Flocon
		53	Vente	14	6	Flocon
		59	Vente	8	2	Flocon
		60	Vente	6	2	Flocon
		62	Décision d'achat	10	4	Flocon
		65	Vente	22	11	Flocon
		68	Vente	3	0	Etoile
		72	Vente	4	2	Flocon
		74	Vente	4	0	Etoile
		78	Vente	4	0	Etoile
		79	Vente	17	5	Flocon

		81	Commande	8	0	Galaxie
		83	Vente	4	0	Etoile
		84	Vente	12	4	Flocon
		85	Vente	21	11	Flocon
		86	Vente	5	0	Etoile
		87	Vente	4	0	Etoile
		88	Vente	11	4	Flocon
		89	Vente	7	0	Etoile
		90	Vente	5	0	Etoile
		91	Vente	5	0	Etoile
		92	Vente	12	6	Flocon
		93	Vente	5	0	Etoile
		94	Vente	7	2	Flocon
		96	Vente	11	4	Flocon
		97	Vente	3	0	Etoile
		98	Vente	3	0	Etoile
		99	Chiffre d'affaires	4	0	Etoile
		100	Chiffre d'affaires	7	3	Flocon
Gestion	14	3	Compte client	4	1	Flocon
		15	Comptabilité	4	0	Etoile
		16	Gestion budget	4	0	Etoile
		17	Transaction client	4	0	Etoile
		18	Gestion employés	5	0	Etoile
		19	Facturation	5	0	Etoile
		25	Compte client	13	5	Flocon
		31	Gestion activité	4	1	Flocon
		40	Epargne	4	0	Etoile
		46	Test qualité	7	0	Etoile
		52	Remboursement	20	4	Flocon

		54	Evènement	10	5	Flocon
		70	Compte client	19	8	Flocon
		80	Transaction	16	8	Flocon
Télécommunications	8	23	Flux de clics	9	0	Etoile
		29	Appel	10	5	Flocon
		45	Trafic internet	5	3	Flocon
		57	Gestion d'emails	5	0	Etoile
		69	Réseaux sociaux	24	12	Flocon
		71	Flux de clics	5	0	Flocon
		81	Flux de clics	9	0	Galaxie
		82	Connexion	5	0	Etoile
Société	8	6	Accident	11	4	Flocon
		7	Maintenance des routes	7	2	Flocon
		20	Vol avion	9	0	Etoile
		21	Réservation hôtel	4	0	Etoile
		24	Assurance	10	0	Etoile
		34	Ticket avion	11	4	Flocon
		55	Accident	16	11	Flocon
		77	Transport public	4	0	Etoile
Education	7	22	Enregistrement étudiant	4	0	Etoile
		33	Inscription étudiant	9	3	Flocon
		39	Evaluation	19	3	Galaxie
		39	Inscription	19	3	Galaxie
		39	Etude	19	3	Galaxie
		43	Etudiants	6	0	Flocon

		66	Facultés universitaires	4	0	Etoile
		67	Cotation université	3	0	Etoile
		73	Examen université	6	1	Flocon
Médical	7	28	Soins animaux	11	2	Flocon
		35	Admissions hôpital	8	4	Flocon
		48	Chirurgie	10	0	Etoile
		49	Chirurgie	9	0	Etoile
		50	Chirurgie	9	0	Etoile
		63	Soins	9	0	Etoile
		64	Médical	4	0	Etoile
Culture	6	36	Critique film	9	1	Flocon
		38	Evaluation de publications	8	4	Flocon
		42	Article	4	4	Flocon
		51	Film	9	1	Flocon
		56	Document	5	2	Flocon
		58	Article journal	3	0	Etoile
Géologie	4	41	Erosion	15	0	Etoile
		47	Erosion	16	0	Etoile
		75	Pollution eau	10	3	Flocon
		76	Production agricole	8	3	Flocon
Autre	2	61	Naissance	7	0	Etoile
		95	Jeu	9	1	Flocon

FIGURE 1: Graphique du nombre de schémas par domaines

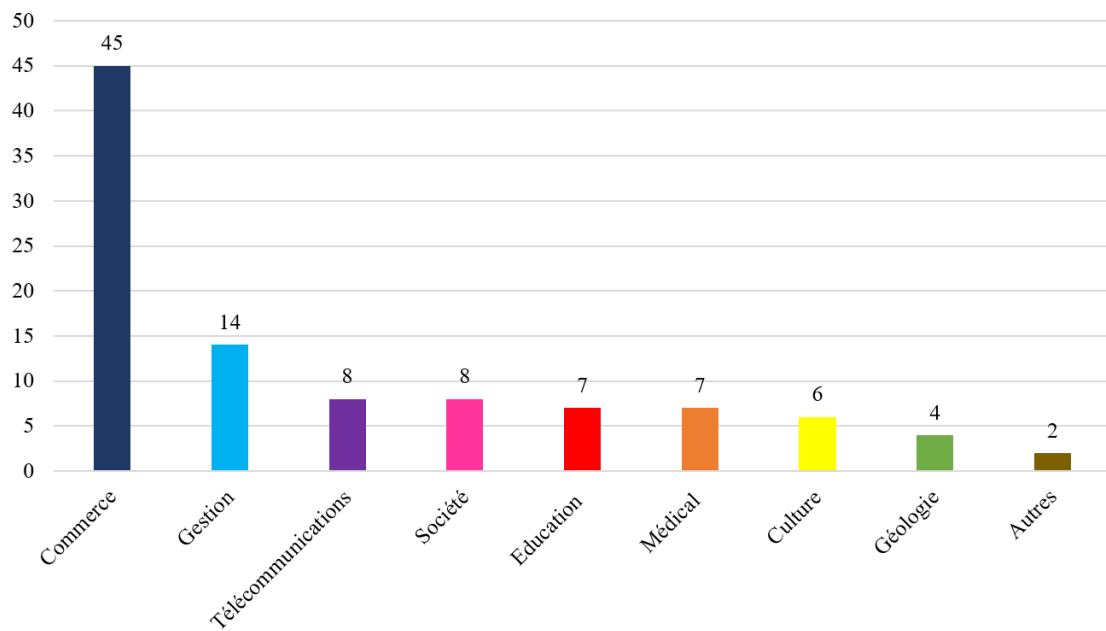


FIGURE 2: Graphique du nombre de schémas par formes par domaines

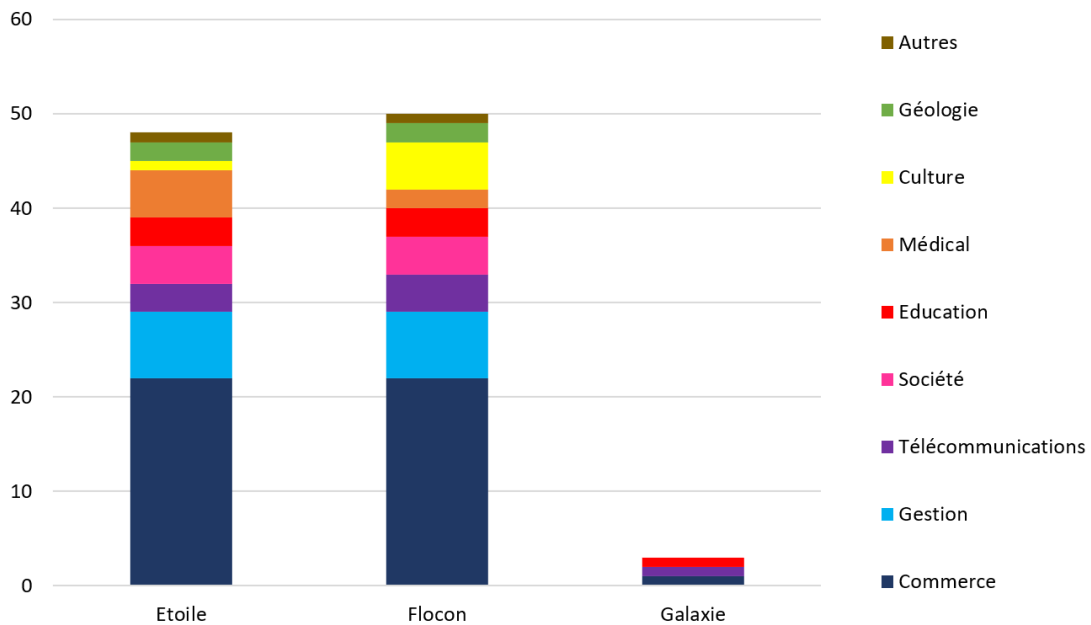
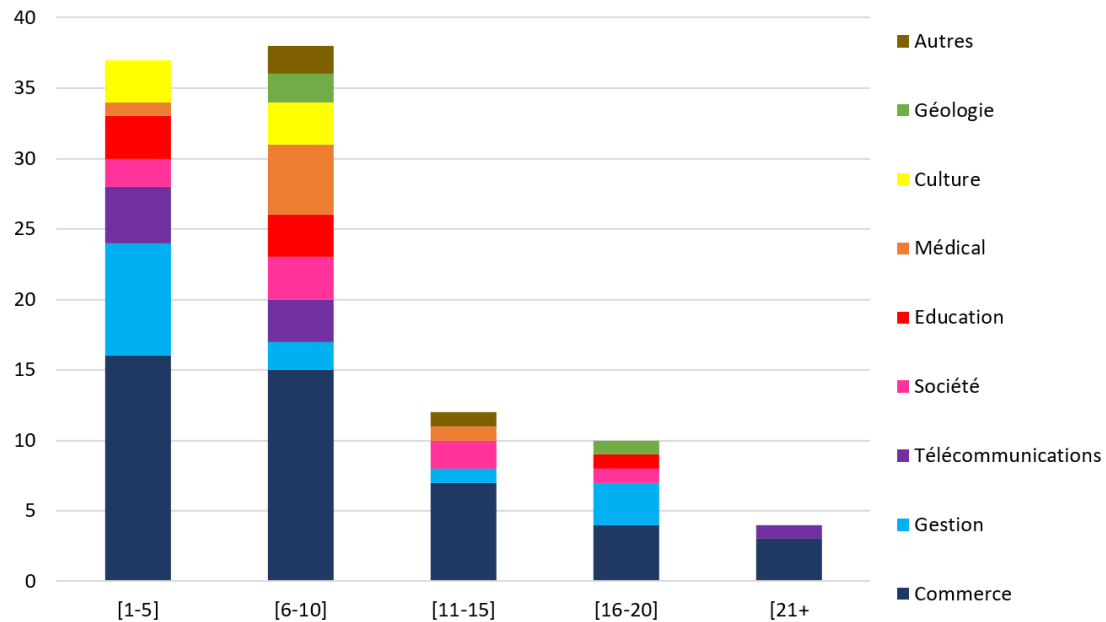


FIGURE 3: Graphique du nombre de schémas par nombre de dimensions par domaines



### 3.3 Application du concept de patterns de design

Après la collecte et l'analyse des schémas, la troisième étape de notre étude consiste à explorer les possibilités d'application du concept de design patterns. Dans un premier temps, cette section analyse quels éléments peuvent servir de base à des potentiels design patterns. Contrairement à l'application du concept dans d'autres domaines tel que celui de l'ingénierie logicielle, ces éléments ne peuvent pas être transversaux. Cette section explique alors comment le concept sera appliqué au cas de la conception de Data Warehouse en évoluant vers des design patterns de domaine. Pour rendre cette adaptation plus riche, chaque design pattern de domaine sera complété par un listing non-exhaustif des exigences auxquelles il permet de répondre. De plus, une matrice de traduction des exigences sera également élaborée.

### 3.3.1 Adaptation du concept

La première intuition pour identifier les éléments sur lesquels les design patterns peuvent se baser est de s'inspirer de ceux de l'ingénierie logicielle. Ils sont basés sur des éléments transversaux à toutes les conceptions. Cependant, l'analyse des schémas de conception de Data Warehouses a permis de mettre en lumière que les seuls éléments transversaux à toutes les conceptions collectées sont leurs éléments structurels et leurs formes. Ces éléments étant simplement les caractéristiques de Data Warehouses, cela ne fait donc pas sens de baser des design patterns sur ceux-ci. Comme expliqué dans le chapitre suivant, on peut émettre l'hypothèse que l'absence de tels éléments vient de la collecte de données et qu'une collecte de données alternative, par exemple en entreprise, pourrait permettre d'identifier des éléments transversaux.

Sur base des conceptions collectées, l'approche idéale de baser des design patterns sur des éléments transversaux à toutes les conceptions tout comme en ingénierie logicielle est donc difficilement accessible. En reprenant la définition d'un design pattern qui est d'être un modèle pouvant être employé dans le cadre d'une problématique bien précise pour réutiliser des solutions établies, une autre approche consiste à identifier des problématiques dans lesquelles il existe des éléments transversaux à toutes les conceptions. L'analyse de la section précédente a permis de mettre en avant huit domaines métier dans lesquels il est possible de regrouper la plupart des conceptions recoltées. Pour le cas de la conception de Data Warehouse, l'utilisation du concept de design pattern peut alors se baser sur ces domaines métier. L'idée est toujours de se baser sur des solutions éprouvées et le but est d'aider les développeurs à la conception. Pour éviter un risque de confusion au concept de design pattern tel qu'utilisé en ingénierie logicielle, nous désignons la contribution ainsi définie et présentée ci-dessous par l'expression design pattern de domaine.

Cette idée d'adapter le concept est confortée par l'analyse des conceptions d'un même domaine. Tout d'abord, on peut constater que les conceptions d'un même domaine métier

ont de nombreuses dimensions communes. On peut en conclure que le sujet d'analyse d'un Data Warehouse est un élément central influençant sa conception. De plus, on constate que 98 des 100 cas de conception collectés peuvent être regroupés dans les huit domaines métier identifiés. On peut donc en conclure que les sujets d'analyse de Data Warehouses tournent souvent autour de thèmes semblables.

Cette adaptation du concept design pattern de domaine est celle qui a le plus de sens pour la conception de Data Warehouses. Différente de l'approche adoptée dans le domaine de l'ingénierie logicielle, celle-ci est un juste milieu entre un design purement conceptuel et un exemple précis de conception. Le nombre limité de patterns structuraux dans les Data Warehouses ouvre la porte à une exploration plus contextuelle de la notion de pattern. De plus, dans l'optique d'aider au mieux les développeurs de Data Warehouses, une adaptation des design patterns vers une utilisation plus concrète semble judicieuse.

### **3.3.2 Lien avec l'identification des exigences**

Dans le but de rendre ces domaines design patterns plus riches et faciles d'utilisation, ils disposent tous d'une liste non exhaustive des exigences auxquelles ils permettent de répondre. Cette liste est non exhaustive pour deux raisons. La première est que les exigences sur un Data Warehouse relèvent d'une stratégie spécifique à chaque organisation. Il est impossible de collecter toutes les exigences possibles et imaginables pour un domaine métier. La deuxième raison est que cette liste a été élaborée à l'aide des informations présentes dans les différentes sources des cas de conception ainsi qu'avec mes connaissances issues de ma formation en ingénieur de gestion. Dans le chapitre suivant plusieurs pistes sont envisagées pour étoffer cette liste.

Ces design patterns de domaine et leur liste d'exigences sont également accompagnés d'une matrice de traduction des exigences. Cette matrice s'inspire du cadré général pour la spécification du besoin d'Esteban Zimanyi et Alejandro Vaisman [2]. Cette matrice



est constituée d'un côté des mesures par faits d'analyse et des différentes dimensions et d'un autre côté des exigences listées. Le but de cette matrice est de faciliter l'utilisation des design patterns de domaine pour les développeurs.

### **3.3.3 Identification des design patterns de domaine**

La démarche d'identification des design patterns de domaine ainsi que de leurs listes d'exigences et matrices de traduction est constituée de six étapes. C'est en partant de l'identification des dimensions que les exigences sont établies. On peut donc considérer que cette démarche est une activité de rétro-ingénierie [27]. La première étape est de lire les sources des cas collectés. La deuxième étape est de réaliser une liste des sujets d'analyse par domaine. Ces deux premières étapes permettent de se mettre dans le contexte de chaque domaine. La troisième étape consiste à analyser les schémas pour identifier les dimensions les plus fréquentes et les plus importantes pour le domaine en question. La quatrième étape est de représenter le design pattern de chaque domaine en reprenant les dimensions identifiées ainsi que la liste des sujets d'analyse potentiels. La cinquième étape consiste à déterminer à quelles exigences les dimensions identifiées permettent de répondre et d'en faire la liste. La sixième et dernière étape est de réaliser une matrice liant les mesures, les dimensions et les exigences auxquelles elles correspondent. Il est important de réaliser ces étapes dans l'ordre et de faire un domaine à la fois.<sup>7</sup>

## **3.4 Design pattern de domaine**

Ce section présente les design patterns de domaine résultant de notre analyse.<sup>8</sup>

---

7. Selon l'analyse des schémas, les schémas ne portant pas sur des thèmes semblables sont regroupés dans le domaine "autre". Il est impossible d'identifier un design pattern de domaine car il ne porte justement pas sur une problématique bien précise.

8. Chaque sous-section est dédiée à un domaine et est organisée selon l'ordre pratiqué lors de la conception de Data Warehouses avec d'abord la liste des exigences, ensuite la matrice de traduction des exigences et enfin le design pattern du domaine.

Les domaines design patterns sont organisés comme des Data Warehouses. Les différentes tables de fait représentent les faits d'analyse possibles par domaine. Dans ces tables se trouvent leurs mesures. Ces faits correspondent aux divers événements pouvant être observés dans ce domaine. Ils ne correspondent pas intégralement à tous les faits des schémas collectés car on constate que ceux-ci ne sont pas toujours réellement des événements.<sup>9</sup> A côté de ces tables de fait, les potentielles dimensions du domaine sont disponibles. Tout comme pour les faits, elles ne correspondent pas intégralement à toutes les dimensions des schémas collectés. Un tri a été réalisé pour ne garder que les plus courantes et pertinentes pour chaque domaine.

### 3.4.1 Domaine du commerce

#### Liste des exigences

1. Analyser les achats :

(a) En terme de quantité :

- i. Comparer la quantité d'achats par dates (années, mois, jours, ...)
- ii. Comparer la quantité d'achats par employés (Marie Dupuis, Jacques Dupond, ...)
- iii. Comparer la quantité d'achats par entreprises (distributeurs, expéditeurs, ...)
- iv. Comparer la quantité d'achats par localisations (pays, région, province, ville, magasin, ...)
- v. Comparer la quantité d'achats par méthodes de paiement (liquide, carte, ...)
- vi. Comparer la quantité d'achats par produits (tables, bananes, smartphone, ...)
- vii. Comparer la quantité d'achats par promotions (soldes, black friday, ...)
- viii. Comparer la quantité d'achats par types (en ligne, en magasin, ...)

(b) En terme de prix :

---

9. Par exemple, dans le domaine du commerce, un client est plus une dimension qu'un événement.

- i. Comparer le prix des achats par dates (années, mois, jours, ...)
- ii. Comparer le prix des achats par employés (Marie Dupuis, Jacques Dupond, ...)
- iii. Comparer le prix des achats par entreprises (Carrefour, Colruyt, ...)
- iv. Comparer le prix des achats par localisations (pays, région, province, ville, magasin, ...)
- v. Comparer le prix des achats par méthodes de paiement (liquide, carte, ...)
- vi. Comparer le prix des achats par produits (tables, bananes, smartphone, ...)
- vii. Comparer le prix des achats par promotions (soldes, black friday, ...)
- viii. Comparer le prix des achats par types (en ligne, en magasin, ...)

## 2. Analyser les exportations :

### (a) En terme de nombre :

- i. Comparer le nombre d'exportations par dates (années, mois, ...)
- ii. Comparer le nombre d'exportations par employés (Marie Dupuis, Jacques Dupond, ...)
- iii. Comparer le nombre d'exportations par entreprises (Carrefour, Colruyt, ...)
- iv. Comparer le nombre d'exportations par localisations (pays, région, province, ville, magasin, ...)
- v. Comparer le nombre d'exportations par produits (tables, bananes, smartphone, ...)
- vi. Comparer le nombre d'exportations par types (camion, bateau, ...)

### (b) En terme de coût :

- i. Comparer le coût des exportations par dates (années, mois, ...)
- ii. Comparer le coût des exportations par employés (Marie Dupuis, Jacques Dupond, ...)

- iii. Comparer le coût des exportations par entreprises (Carrefour, Colruyt, ...)
- iv. Comparer le coût des exportations par localisations (pays, région, province, ville, magasin, ...)
- v. Comparer le coût des exportations par produits (tables, bananes, smartphone, ...)
- vi. Comparer le coût des exportations par types (camion, bateau, ...)

### 3. Analyser les inventaires :

#### (a) En terme de nombre :

- i. Analyser le nombre d'inventaires par dates (année, mois, ...)
- ii. Analyser le nombre d'inventaires par employés (Marie Dupuis, Jacques Dupond, ...)
- iii. Analyser le nombre d'inventaires par localisations (pays, ville, entrepôt, ...)
- iv. Analyser le nombre d'inventaires par produits (tables, bananes, smartphone, ...)

#### (b) En terme de taille :

- i. Comparer la taille des inventaires par dates (année, mois, ...)
- ii. Comparer la taille des inventaires par employés (Marie Dupuis, Jacques Dupond, ...)
- iii. Comparer la taille des inventaires par localisations (pays, ville, entrepôt, ...)
- iv. Comparer la taille des inventaires par produits (tables, bananes, smartphone, ...)

### 4. Analyser les ventes :

#### (a) En terme de quantité :

- i. Comparer les quantités vendues par clients (Paul Dupuis, Sophie Dubois, ...)
- ii. Comparer les quantités vendues par dates (années, mois, ...)

- iii. Comparer les quantités vendues par employés (Marie Dupuis, Jacques Dupond, ...)
  - iv. Comparer les quantités vendues par localisations (pays, région, magasins, ...)
  - v. Comparer les quantités vendues par modes de paiement (en liquide, par carte, ...)
  - vi. Comparer les quantités vendues par produits (tables, bananes, smartphone, ...)
  - vii. Comparer les quantités vendues par promotions (soldes, black friday, ...)
  - viii. Comparer les quantités vendues par types de ventes (en ligne, en magasin, ...)
- (b) En terme de prix :
- i. Comparer le prix des ventes par clients (Paul Dupuis, Sophie Dubois, ...)
  - ii. Comparer le prix des ventes par dates (années, mois, ...)
  - iii. Comparer le prix des ventes par employés (Marie Dupuis, Jacques Dupond, ...)
  - iv. Comparer le prix des ventes par localisations (pays, région, magasins, ...)
  - v. Comparer le prix des ventes par modes de paiement (en liquide, par carte, ...)
  - vi. Comparer le prix des ventes par produits (tables, bananes, smartphone, ...)
  - vii. Comparer le prix des ventes par promotions (soldes, black friday, ...)
  - viii. Comparer le prix des ventes par types de ventes (en ligne, en magasin, ...)
- (c) En terme de profit :
- i. Comparer le profit des ventes par clients (Paul Dupuis, Sophie Dubois, ...)
  - ii. Comparer le profit des ventes par dates (années, mois, ...)

- iii. Comparer le profit des ventes par employés (Marie Dupuis, Jacques Dupond, ...)
- iv. Comparer le profit des ventes par localisations (pays, région, magasins, ...)
- v. Comparer le profit des ventes par modes de paiement (en liquide, par carte, ...)
- vi. Comparer le profit des ventes par produits (tables, bananes, smartphone, ...)
- vii. Comparer le profit des ventes par promotions (soldes, black friday, ...)
- viii. Comparer le profit des ventes par types de ventes (en ligne, en magasin, ...)

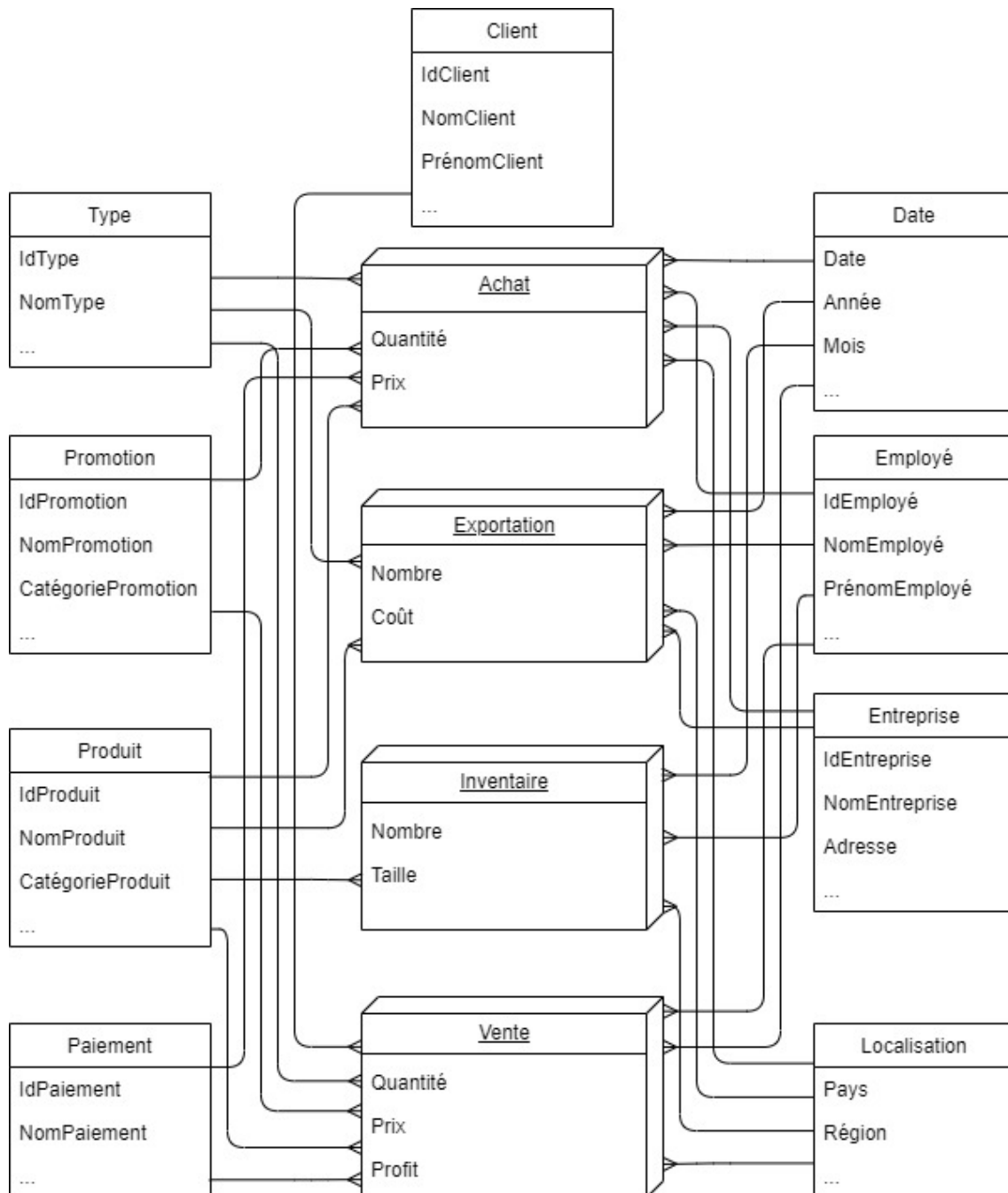
### Matrice de traduction des exigences

TABLE 2: Matrice de traduction des exigences du domaine du commerce

		Exigences								
		1a	1b	2a	2b	3a	3b	4a	4b	4c
Faits et mesures	Achat	✓	✓							
	Quantité	✓								
	Prix		✓							
	Exportation			✓	✓					
	Nombre			✓						
	Coût				✓					
	Inventaire					✓	✓			
	Nombre					✓				
	Taille						✓			
	Vente							✓	✓	✓
	Quantité							✓		
	Prix								✓	
	Profit									✓
Dimensions	Client							✓	✓	✓
	Date	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Employé	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Entreprise	✓	✓	✓	✓					
	Localisation	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Paiement	✓	✓					✓	✓	✓
	Produit	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Promotion	✓	✓					✓	✓	✓
	Type	✓	✓	✓	✓			✓	✓	✓

## Design pattern de domaine

FIGURE 4: Design pattern du domaine du commerce



### 3.4.2 Domaine de la gestion

#### Liste des exigences

##### 1. Analyser les activités :

###### (a) En terme de nombre :

- i. Analyser le nombre d'activités par clients (Paul Dupuis, Sophie Dubois, ...)
- ii. Analyser le nombre d'activités par dates (année, mois, ...)
- iii. Analyser le nombre d'activités par employés (Marie Dupuis, Jacques Dupond, ...)
- iv. Analyser le nombre d'activités par localisations (pays, région, ...)
- v. Analyser le nombre d'activités par types (activités d'entreprises, événements, ...)

###### (b) En terme de durée :

- i. Analyser la durée des activités par clients (Paul Dupuis, Sophie Dubois, ...)
- ii. Analyser la durée des activités par dates (année, mois, ...)
- iii. Analyser la durée des activités par employés (Marie Dupuis, Jacques Dupond, ...)
- iv. Analyser la durée des activités par localisations (pays, région, ...)
- v. Analyser la durée des activités par types (activités d'entreprises, événements, ...)

##### 2. Analyser la facturation :

###### (a) En terme de taxes :

- i. Analyser les taxes des factures par clients (Paul Dupuis, Sophie Dubois, ...)
- ii. Analyser les taxes des factures par dates (année, mois, ...)
- iii. Analyser les taxes des factures par entreprises (Côte d'Or, Leonidas, ...)



- iv. Analyser les taxes des factures par localisations (pays, ...)
  - v. Analyser les taxes des factures par produits (tables, bananes, smartphone, ...)
- (b) En terme de quantité :
- i. Analyser la quantité de factures par clients (Paul Dupuis, Sophie Dubois, ...)
  - ii. Analyser la quantité de factures par dates (année, mois, ...)
  - iii. Analyser la quantité de factures par entreprises (Côte d'Or, Leonidas, ...)
  - iv. Analyser la quantité de factures par localisations (pays, ...)
  - v. Analyser la quantité de factures par produits (tables, bananes, smartphone, ...)
- (c) En terme de montant :
- i. Analyser le montant des factures par clients (Paul Dupuis, Sophie Dubois, ...)
  - ii. Analyser le montant des factures par dates (année, mois, ...)
  - iii. Analyser le montant des factures par entreprises (Côte d'Or, Leonidas, ...)
  - iv. Analyser le montant des factures par localisations (pays, ...)
  - v. Analyser le montant des factures par produits (tables, bananes, smartphone, ...)
3. Analyser les tests de qualité :
- (a) En terme de quantité :
- i. Analyser la quantité de tests de qualité par entreprises (Côte d'Or, Leonidas, ...)
  - ii. Analyser la quantité de tests de qualité par dates (année, mois, ...)
  - iii. Analyser la quantité de tests de qualité par produits (nourriture, object, ...)
  - iv. Analyser la quantité de tests de qualité par types (tests, normes, ...)

(b) En terme de résultats :

- i. Analyser les résultats de tests de qualité par entreprises (Côte d'Or, Leonidas, ...)
- ii. Analyser les résultats de tests de qualité par dates (année, mois, ...)
- iii. Analyser les résultats de tests de qualité par produits (nourriture, objet, ...)
- iv. Analyser les résultats de tests de qualité par types (tests, normes, ...)

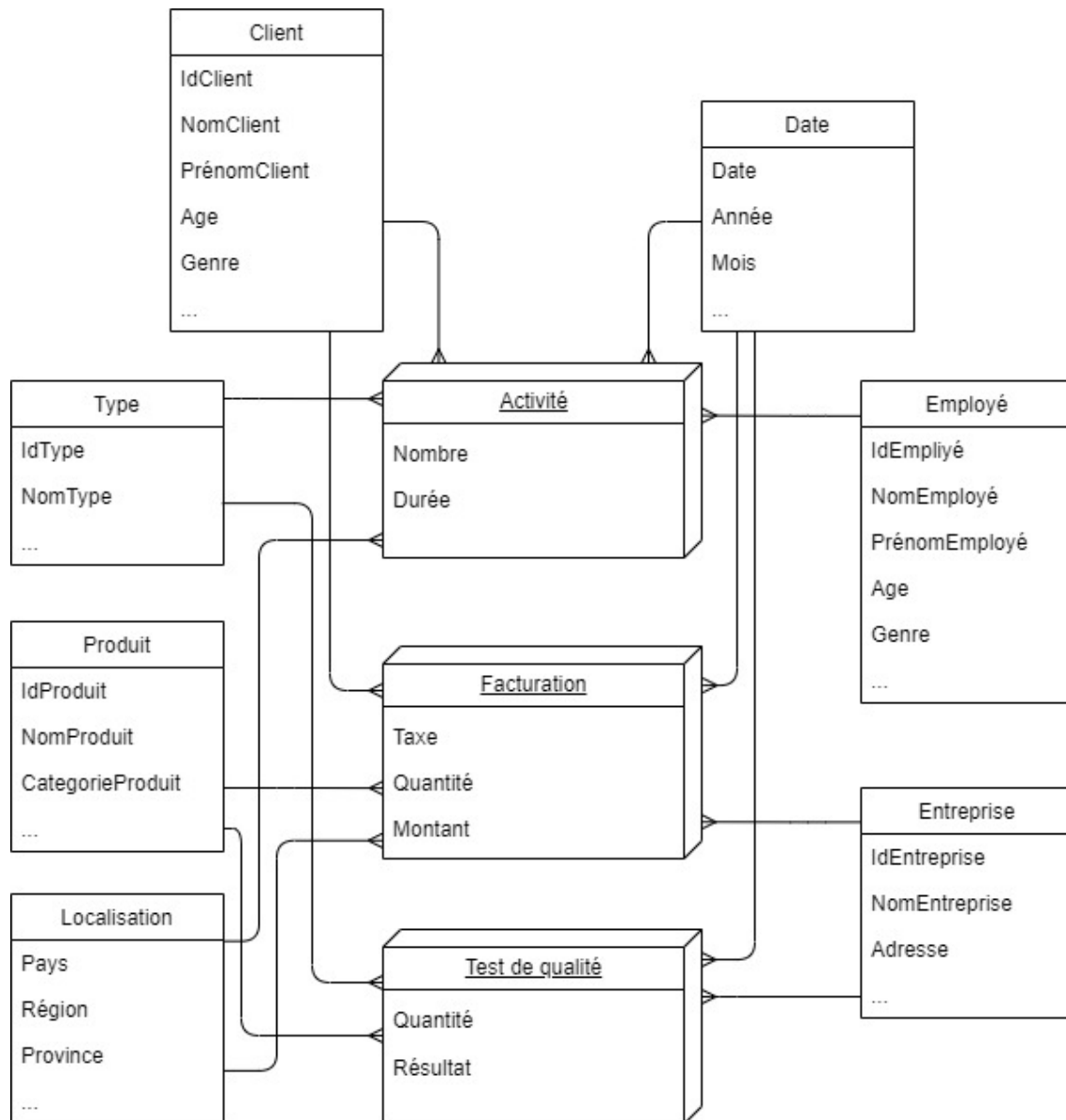
### Matrice de traduction des exigences

TABLE 3: Matrice de traduction des exigences du domaine de la gestion

		Exigences						
		1a	1b	2a	2b	2c	3a	3b
Faits et mesures	Activité	✓	✓					
	Nombre	✓						
	Durée		✓					
	Facturation			✓	✓	✓		
	Taxe			✓				
	Quantité				✓			
	Montant					✓		
	Test de qualité						✓	✓
	Quantité						✓	
	Résultat							✓
Dimensions	Client	✓	✓	✓	✓	✓		
	Date	✓	✓	✓	✓	✓	✓	✓
	Employé	✓	✓					
	Entreprise			✓	✓	✓	✓	✓
	Localisation	✓	✓	✓	✓	✓		
	Produit			✓	✓	✓	✓	✓
	Type	✓	✓				✓	✓

## Design pattern de domaine

FIGURE 5: Design pattern du domaine de la gestion



### 3.4.3 Domaine des télécommunications

#### Liste des exigences

##### 1. Analyser les appels téléphoniques :

###### (a) En terme de nombre :

- i. Analyser le nombre d'appels téléphoniques par dates (année, mois, ...)
- ii. Analyser le nombre d'appels téléphoniques par hôtes destinataires (personne appelée)
- iii. Analyser le nombre d'appels téléphoniques par hôtes sources (personne appelant)
- iv. Analyser le nombre d'appels téléphoniques par localisations (pays, région, province, ...)
- v. Analyser le nombre d'appels téléphoniques par utilisateurs (client analysé)

###### (b) En terme de durée :

- i. Analyser la durée des appels téléphoniques par dates (année, mois, ...)
- ii. Analyser la durée des appels téléphoniques par hôtes destinataires (personne appelée)
- iii. Analyser la durée des appels téléphoniques par hôtes sources (personne appelant)
- iv. Analyser la durée des appels téléphoniques par localisations (pays, région, province, ...)
- v. Analyser la durée des appels téléphoniques par utilisateurs (client analysé)

###### (c) En terme de coût :

- i. Analyser le coût d'appels téléphoniques par dates (année, mois, ...)
- ii. Analyser le coût d'appels téléphoniques par hôtes destinataires (personne appelée)

- iii. Analyser le coût d'appels téléphoniques par hôtes sources (personne appelant)
  - iv. Analyser le coût d'appels téléphoniques par localisations (pays, région, province, ...)
  - v. Analyser le coût d'appels téléphoniques par utilisateurs (client analysé)
2. Analyser les clics :
- (a) En terme de nombre :
    - i. Analyser le nombre de clics par dates (année, mois, ...)
    - ii. Analyser le nombre de clics par localisations (pays, région, ...)
    - iii. Analyser le nombre de clics par pages (Webcampus, Facebook, ...)
    - iv. Analyser le nombre de clics par utilisateurs (Marie Dupuis, Jacques Dupond, ...)
3. Analyser les connexions :
- (a) En terme de quantité :
    - i. Analyser la quantité de connexions par dates (année, mois, ...)
    - ii. Analyser la quantité de connexions par hôtes destinataires (destination de l'adresse IP et port)
    - iii. Analyser la quantité de connexions par hôtes sources (source de l'adresse IP et port)
    - iv. Analyser la quantité de connexions par localisations (pays, région, province, ...)
    - v. Analyser la quantité de connexions par pages (Webcampus, Facebook, ...)
  - (b) En terme de durée :
    - i. Analyser la quantité de connexions par dates (année, mois, ...)
    - ii. Analyser la quantité de connexions par hôtes destinataires (destination de l'adresse IP et port)

- iii. Analyser la quantité de connexions par hôtes sources (source de l'adresse IP et port)
- iv. Analyser la quantité de connexions par localisations (pays, région, province, ...)
- v. Analyser la quantité de connexions par pages (Webcampus, Facebook, ...)

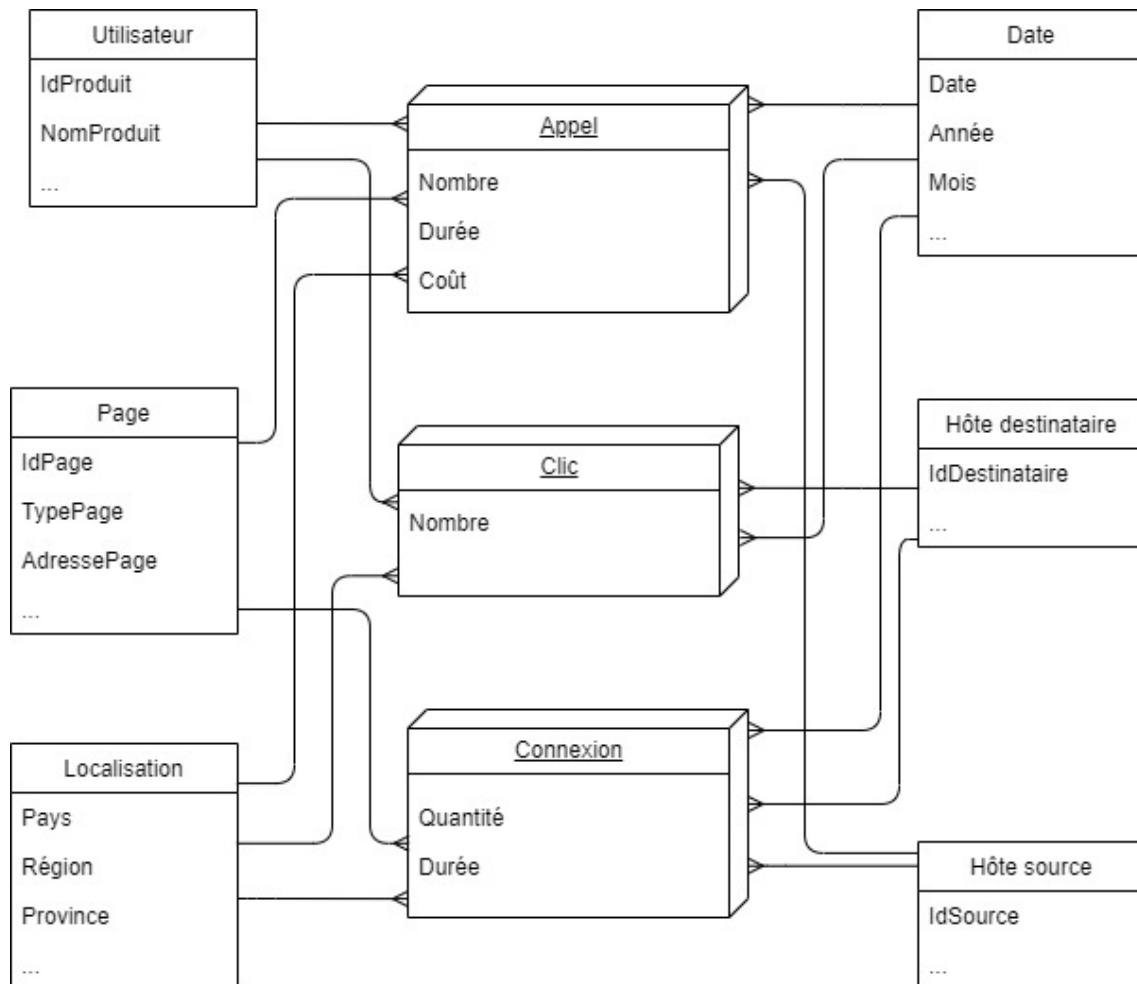
### Matrice de traduction des exigences

TABLE 4: Matrice de traduction des exigences du domaine des télécommunications

		Exigences					
		1a	1b	1c	2a	3a	3b
Faits et mesures	Appel	✓	✓	✓			
	Nombre	✓					
	Durée		✓				
	Coût			✓			
	Clic				✓		
	Nombre				✓		
	Connexion					✓	✓
	Quantité					✓	
	Durée						✓
Dimensions	Date	✓	✓	✓	✓	✓	✓
	Hôte destinataire	✓	✓	✓		✓	✓
	Hôte source	✓	✓	✓		✓	✓
	Localisation	✓	✓	✓	✓	✓	✓
	Page				✓	✓	✓
	Utilisateur	✓	✓	✓	✓		

## Design pattern de domaine

FIGURE 6: Design pattern du domaine des télécommunications



### 3.4.4 Domaine société

#### Liste des exigences

##### 1. Analyser les accidents :

###### (a) En terme de nombre :

- i. Analyser le nombre d'accidents par clients (Marie Dupuis, Marc Dupond, ...)
- ii. Analyser le nombre d'accidents par dates (annnée, mois, ...)
- iii. Analyser le nombre d'accidents par localisations (pays, région, ...)
- iv. Analyser le nombre d'accidents par produits assurés (maison, voiture, ...)
- v. Analyser le nombre d'accidents par types d'assurances (omnium, RC, ...)

###### (b) En terme de coût :

- i. Analyser le coût des accidents par clients (Marie Dupuis, Marc Dupond, ...)
- ii. Analyser le coût des accidents par dates (annnée, mois, ...)
- iii. Analyser le coût des accidents par localisations (pays, région, ...)
- iv. Analyser le coût des accidents par produits assurés (maison, voiture, ...)
- v. Analyser le coût des accidents par types d'assurances (omnium, RC, ...)

##### 2. Analyser la maintenance des routes :

###### (a) En terme de nombre :

- i. Analyser le nombre de maintenances de route par dates (année, mois, ...)
- ii. Analyser le nombre de maintenances de route par localisations (région, province, ...)
- iii. Analyser le nombre de maintenances par routes (E19, N5, ...)

###### (b) En terme de durée :

- i. Analyser la durée de maintenances de route par dates (année, mois, ...)
- ii. Analyser la durée de maintenances de route par localisations (région, province, ...)



- iii. Analyser la durée de maintenances par routes (E19, N5, ...)
- (c) En terme de coût :
  - i. Analyser le coût de maintenances de route par dates (année, mois, ...)
  - ii. Analyser le coût de maintenances de route par localisations (région, province, ...)
  - iii. Analyser le coût de maintenances par routes (E19, N5, ...)
- 3. Analyser les réservations d'hôtel :
  - (a) En terme de nombre :
    - i. Analyser le nombre de réservations d'hôtel par canaux de vente (agence, internet, ...)
    - ii. Analyser le nombre de réservations d'hôtel par clients (Marie Dupond, Jacques Dupuis, ...)
    - iii. Analyser le nombre de réservations d'hôtel par dates (année, mois, ...)
    - iv. Analyser le nombre de réservations d'hôtel par localisations de départ/réservation (pays, région, ...)
    - v. Analyser le nombre de réservations d'hôtel par destinations (pays, région, ...)
    - vi. Analyser le nombre de réservations d'hôtel par hôtels (hôtel de la poste, hôtel de la plage, ...)
  - (b) En terme de durée :
    - i. Analyser la durée du séjour des réservations d'hôtel par canaux de vente (agence, internet, ...)
    - ii. Analyser la durée du séjour des réservations d'hôtel par clients (Marie Dupond, Jacques Dupuis, ...)
    - iii. Analyser la durée du séjour des réservations d'hôtel par dates (année, mois, ...)

- iv. Analyser la durée du séjour des réservations d'hôtel par localisations de départ/réservation (pays, région, ...)
  - v. Analyser la durée du séjour des réservations d'hôtel par destinations (pays, région, ...)
  - vi. Analyser la durée du séjour des réservations d'hôtel par hôtels (hôtel de la poste, hôtel de la plage, ...)
- (c) En terme de prix :
- i. Analyser le prix des réservations d'hôtel par canaux de vente (agence, internet, ...)
  - ii. Analyser le prix des réservations d'hôtel par clients (Marie Dupond, Jacques Dupuis, ...)
  - iii. Analyser le prix des réservations d'hôtel par dates (année, mois, ...)
  - iv. Analyser le prix des réservations d'hôtel par localisations de départ/réservation (pays, région, ...)
  - v. Analyser le prix des réservations d'hôtel par destinations (pays, région, ...)
  - vi. Analyser le prix des réservations d'hôtel par hôtels (hôtel de la poste, hôtel de la plage, ...)
4. Analyser les vols d'avion :
- (a) En terme de nombre :
- i. Analyser le nombre de vols d'avions par avions (Boeing 737, airbus a300, ...)
  - ii. Analyser le nombre de vols d'avions par clients (Marie Dupuis, Marc Dupond, ...)
  - iii. Analyser le nombre de vols d'avions par dates (annnée, mois, ...)
  - iv. Analyser le coût de vols d'avions par aéroports de départ (Charleroi, Bruxelles, ...)

- v. Analyser le coût de vols d'avions par aéroports de destinations (Charleroi, Bruxelles, ...)

(b) En terme de coût :

- i. Analyser le coût de vols d'avions par avions (Boeing 737, airbus a300, ...)
- ii. Analyser le coût de vols d'avions par clients (Marie Dupuis, Marc Dupond, ...)
- iii. Analyser le coût de vols d'avions par dates (année, mois, ...)
- iv. Analyser le coût de vols d'avions par aéroports de départ (Charleroi, Bruxelles, ...)
- v. Analyser le coût de vols d'avions par aéroports de destinations (Charleroi, Bruxelles, ...)

(c) En terme de prix :

- i. Analyser le prix de vols d'avions par avions (Boeing 737, airbus a300, ...)
- ii. Analyser le prix de vols d'avions par clients (Marie Dupuis, Marc Dupond, ...)
- iii. Analyser le prix de vols d'avions par dates (année, mois, ...)
- iv. Analyser le prix de vols d'avions par aéroports de départ (Charleroi, Bruxelles, ...)
- v. Analyser le prix de vols d'avions par aéroports de destinations (Charleroi, Bruxelles, ...)

(d) En terme de durée :

- i. Analyser la durée de vols d'avions par avions (Boeing 737, airbus a300, ...)
- ii. Analyser la durée de vols d'avions par clients (Marie Dupuis, Marc Dupond, ...)
- iii. Analyser la durée de vols d'avions par dates (année, mois, ...)

- iv. Analyser la durée de vols d'avions par aéroports de départ (Charleroi, Bruxelles, ...)
- v. Analyser la durée de vols d'avions par aéroports de destinations (Charleroi, Bruxelles, ...)

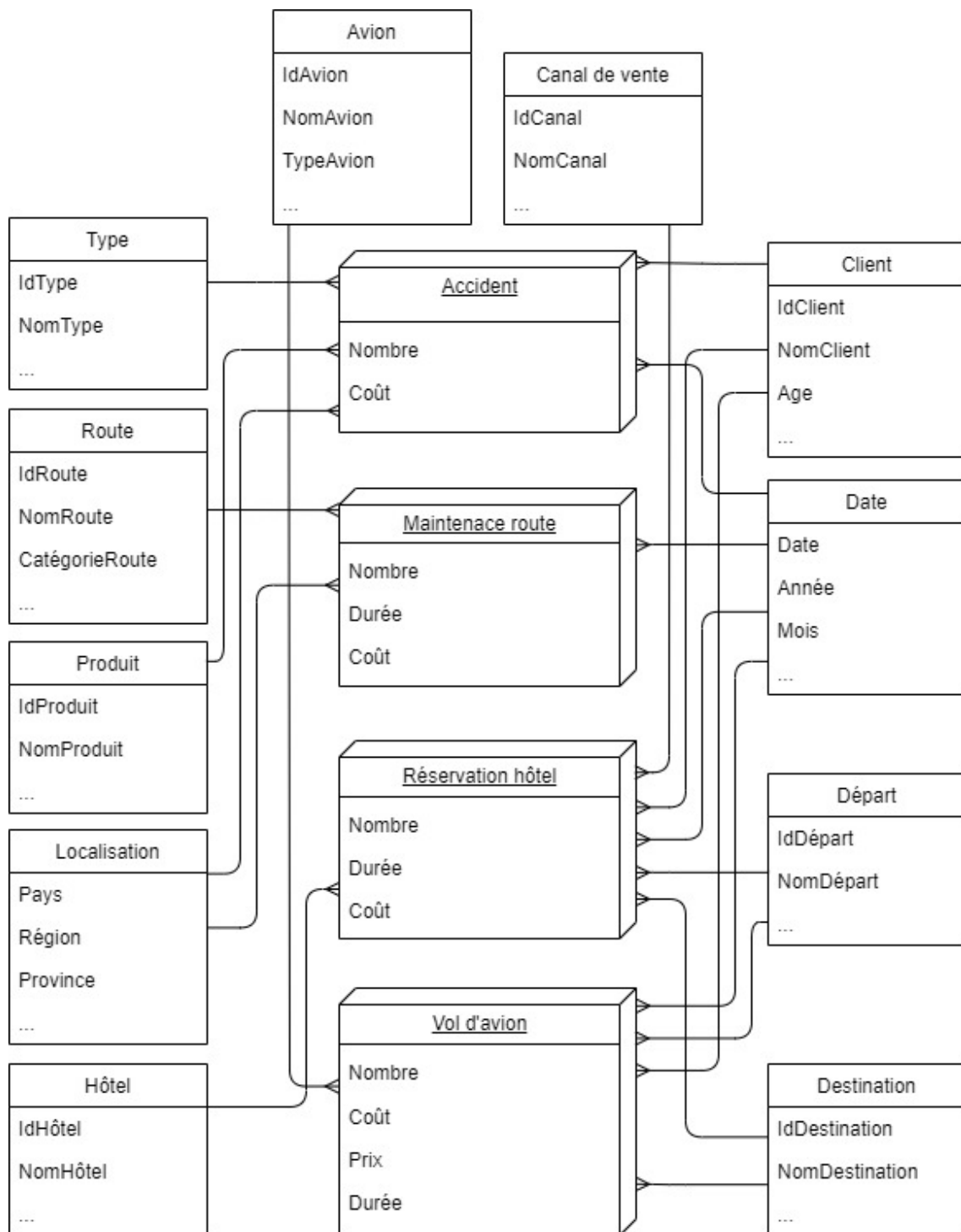
### Matrice de traduction des exigences

TABLE 5: Matrice de traduction des exigences du domaine société

		Exigences											
		1a	1b	2a	2b	2c	3a	3b	3c	4a	4b	4c	4d
Faits et mesures	Accident	✓	✓										
	Nombre	✓											
	Coût		✓										
	Route			✓	✓	✓							
	Nombre			✓									
	Durée				✓								
	Coût					✓							
	Hôtel						✓	✓	✓				
	Nombre						✓						
	Durée							✓					
	Prix								✓				
	Avion									✓	✓	✓	✓
	Nombre									✓			
	Coût										✓		
Dimensions	Prix											✓	
	Durée												✓
	Avion									✓	✓	✓	✓
	Canal de vente						✓	✓	✓				
	Client	✓	✓				✓	✓	✓	✓	✓	✓	✓
	Date	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Départ						✓	✓	✓	✓	✓	✓	✓
	Destination						✓	✓	✓	✓	✓	✓	✓
	Hôtel						✓	✓	✓				
	Localisation	✓	✓	✓	✓	✓							
	Produit	✓	✓										
	Route			✓	✓	✓							
	Type	✓	✓										

## Design pattern de domaine

FIGURE 7: Design pattern du domaine société



### 3.4.5 Domaine de l'éducation

#### Liste des exigences

##### 1. Analyser les évaluations :

###### (a) En terme de nombre :

- i. Analyser le nombre d'évaluations par années d'études (bac1, master2, ...)
- ii. Analyser le nombre d'évaluations par cours (Business Intelligence, Anglais, ...)
- iii. Analyser le nombre d'évaluations par dates (année, mois, ...)
- iv. Analyser le nombre d'évaluations par facultés (médecine, droit, ...)
- v. Analyser le nombre d'évaluations par établissements précédents des étudiants (collège du Sacré-Coeur, institut Sainte-Marie, ...)
- vi. Analyser le nombre d'évaluations par études (ingénieur de gestion, pharmacie, ...)
- vii. Analyser le nombre d'évaluations par étudiants (Pierre Dupond, Jacques Dupuis, ...)
- viii. Analyser le nombre d'évaluations par professeurs (Mme. Durant, M. Dupond, ...)

###### (b) En terme de note :

- i. Analyser les notes des évaluations par années d'études (bac1, master2, ...)
- ii. Analyser les notes des évaluations par cours (Business Intelligence, Anglais, ...)
- iii. Analyser les notes des évaluations par dates (année, mois, ...)
- iv. Analyser les notes des évaluations par facultés (médecine, droit, ...)
- v. Analyser les notes des évaluations par établissements précédents des étudiants (collège du Sacré-Coeur, institut Sainte-Marie, ...)

- vi. Analyser les notes des évaluations par études (ingénieur de gestion, pharmacie, ...)
- vii. Analyser les notes des évaluations par étudiants (Pierre Dupond, Jacques Dupuis, ...)
- viii. Analyser les notes des évaluations par professeurs (Mme. Durant, M. Dupond, ...)

2. Analyser les inscriptions :

(a) En terme de nombre :

- i. Analyser le nombre d'inscriptions par années d'études (bac1, master2, ...)
- ii. Analyser le nombre d'inscriptions par cours (Business Intelligence, Anglais, ...)
- iii. Analyser le nombre d'inscriptions par dates (année, mois, ...)
- iv. Analyser le nombre d'inscriptions par facultés (médecine, droit, ...)
- v. Analyser le nombre d'inscriptions par établissements précédents des étudiants (collège du Sacré-Coeur, institut Sainte-Marie, ...)
- vi. Analyser le nombre d'inscriptions par études (ingénieur de gestion, pharmacie, ...)
- vii. Analyser le nombre d'inscriptions par étudiants (Pierre Dupond, Jacques Dupuis, ...)

(b) En terme de coût :

- i. Analyser les coûts d'inscriptions par années d'études (bac1, master2, ...)
- ii. Analyser les coûts d'inscriptions par cours (Business Intelligence, Anglais, ...)
- iii. Analyser les coûts d'inscriptions par dates (année, mois, ...)
- iv. Analyser les coûts d'inscriptions par facultés (médecine, droit, ...)

- v. Analyser les coûts d'inscriptions par établissements précédents des étudiants (collège du Sacré-Coeur, institut Sainte-Marie, ...)
- vi. Analyser les coûts d'inscriptions par études (ingénieur de gestion, pharmacie, ...)
- vii. Analyser les coûts d'inscriptions par étudiants (Pierre Dupond, Jacques Dupuis, ...)

### Matrice de traduction des exigences

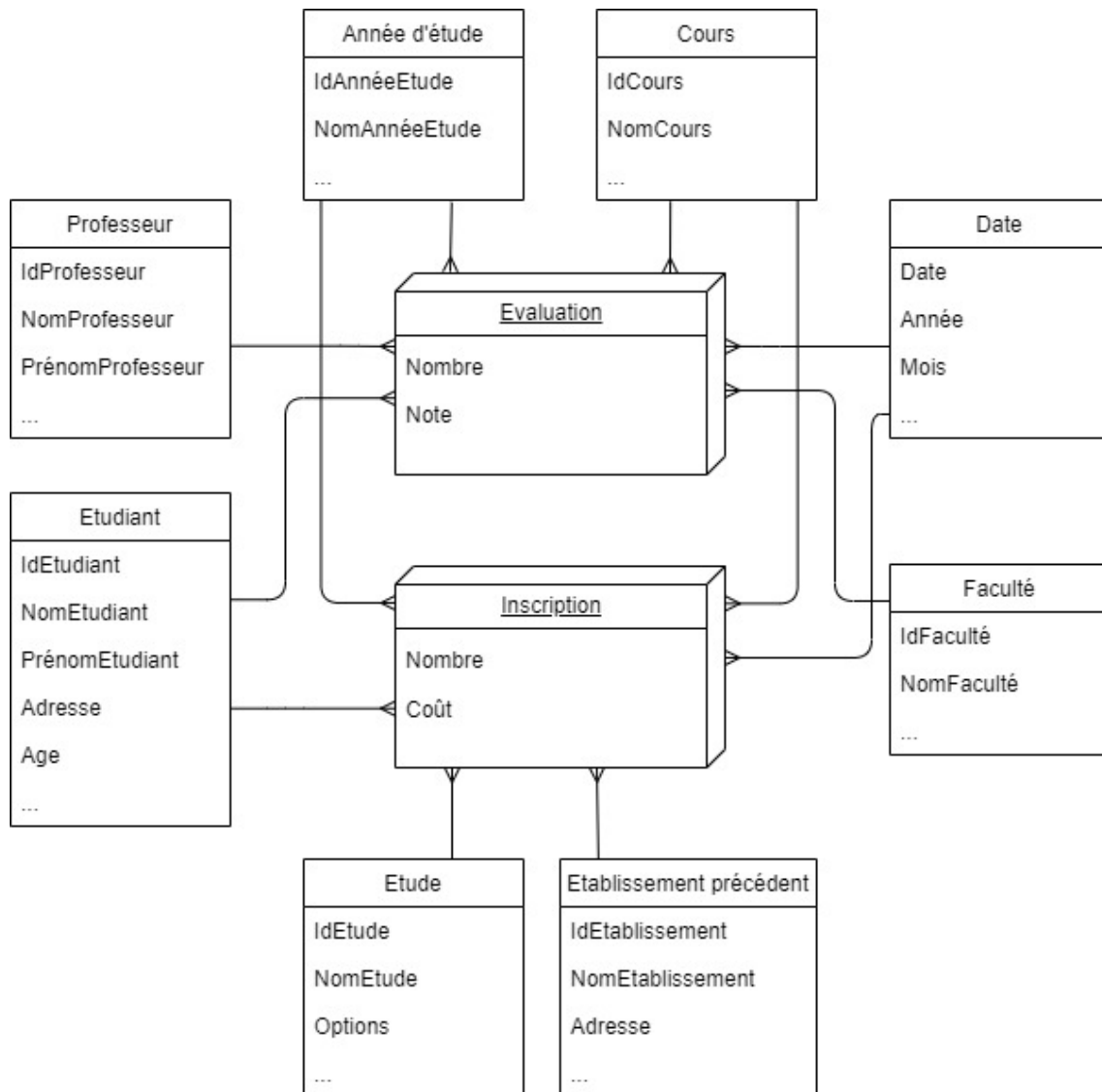
TABLE 6: Matrice de traduction des exigences du domaine de l'éducation

		Exigences			
		1a	1b	2a	2b
Faits et mesures	Evaluation	✓	✓		
	Nombre	✓			
	Note		✓		
	Inscription			✓	✓
	Nombre			✓	
	Coût				✓
Dimensions	Année d'étude	✓	✓	✓	✓
	Cours	✓	✓	✓	✓
	Dates	✓	✓	✓	✓
	Faculté	✓	✓	✓	✓
	Etablissement précédent	✓	✓	✓	✓
	Etude	✓	✓	✓	✓
	Etudiant	✓	✓	✓	✓
	Professeur	✓	✓		



## Design pattern de domaine

FIGURE 8: Design pattern du domaine de l'éducation



### 3.4.6 Domaine médical

#### Liste des exigences

##### 1. Analyser les admissions :

##### (a) En terme de quantité :

- i. Suivre l'évolution du nombre d'admissions selon les dates (années, mois, jours, heures, ...)
- ii. Comparer le nombre d'admissions par départements (cardiologie, chirurgie, urgences, ...)
- iii. Comparer le nombre d'admissions par diagnostics (cancer du sein, problème cardiaque, ...)
- iv. Comparer le nombre d'admissions par hôpitaux (Marie Curie, GHdC, ...)
- v. Comparer le nombre d'admissions par soignants (Docteur Jacques, Infirmier Dupond, ...)
- vi. Comparer le nombre d'admissions par patients (âge, genre, métier, ...)
- vii. Comparer le nombre d'admissions par responsables (Infirmière Dupuis, Docteur Jacques, ...)
- viii. Comparer le nombre d'admissions par traitement (opération du coeur, surveillance des constantes, ... )

##### (b) En terme de coût :

- i. Suivre l'évolution du nombre d'admissions selon les dates (années, mois, jours, heures, ...)
- ii. Comparer le coût des admissions par départements (cardiologie, chirurgie, urgences, ...)
- iii. Comparer le coût des admissions par diagnostics (cancer du sein, problème cardiaque, ...)

- iv. Comparer le coût des admissions par hôpitaux (Marie Curie, GHdC, ...)
- v. Comparer le coût des admissions par soignants (Docteur Jacques, Infirmier Dupond, ...)
- vi. Comparer le coût des admissions par patients (âge, genre, métier, ...)
- vii. Comparer le coût des admissions par responsables (Infirmière chef Mme. Dupond, Dr. Jacques, ...)
- viii. Comparer le coût des admissions par traitements (opération du coeur, surveillance des constantes, ...)

## 2. Analyser les chirurgies :

### (a) En terme de quantité :

- i. Comparer le nombre de chirurgies par date (années, mois, jours, heures, ...)
- ii. Comparer le nombre de chirurgies par diagnostics (cancer du sein, pied cassé, ...)
- iii. Comparer le nombre de chirurgies par hôpitaux (Marie Curie, GHdC, ...)
- iv. Comparer le nombre de chirurgies par soignants (Docteur Jacques, Infirmier Dupond, ...)
- v. Comparer le nombre de chirurgies par patients (âge, genre, métier, ...)
- vi. Comparer le nombre de chirurgies par responsables (Infirmière Dupuis, Docteur Jacques, ...)
- vii. Comparer le nombre de chirurgies par spécialité (neurochirurgie, chirurgie cardiaque, ...)
- viii. Comparer le nombre de chirurgies par zones du corps (jambes, coeur, ...)

### (b) En terme de coût :

- i. Comparer le coût de chirurgies par dates (années, mois, jours, heures, ...)
- ii. Comparer le coût de chirurgies par diagnostics (cancer du sein, pied cassé, ...)

- iii. Comparer le coût de chirurgies par hôpitaux (Marie Curie, GHdC, ...)
  - iv. Comparer le coût de chirurgies par soignants (Docteur Jacques, Infirmier Dupond, ...)
  - v. Comparer le coût de chirurgies par patients (âge, genre, métier, ...)
  - vi. Comparer le coût de chirurgies par responsables (Infirmière Dupuis, Docteur Jacques, ...)
  - vii. Comparer le coût de chirurgies par spécialités (neurochirurgie, chirurgie cardiaque, ...)
  - viii. Comparer le coût de chirurgies par zones du corps (jambes, coeur, ...)
3. Analyser les soins :
- (a) En terme de quantité :
    - i. Comparer le nombre de soins par dates (années, mois, jours, heures, ...)
    - ii. Comparer le nombre de soins par départements (dermatologie, urgences, ...)
    - iii. Comparer le nombre de soins par diagnostics (cancer du sein, pied cassé, ...)
    - iv. Comparer le nombre de soins par hôpitaux (Marie Curie, GHdC, ...)
    - v. Comparer le nombre de soins par soignants (Docteur Jacques, Infirmier Dupond, ...)
    - vi. Comparer le nombre de soins par patients (âge, genre, métier, ...)
    - vii. Comparer le nombre de soins par responsables (Infirmière Dupuis, Docteur Jacques, ...)
    - viii. Comparer le nombre de soins par zones du corps (jambes, coeur, ...)
  - (b) En terme de coût :
    - i. Comparer le coût de soins par dates (années, mois, jours, heures, ...)
    - ii. Comparer le coût de soins par départements (dermatologie, urgences, ...)
    - iii. Comparer le coût de soins par diagnostics (cancer du sein, pied cassé, ...)

- iv. Comparer le coût de soins par hôpitaux (Marie Curie, GHdC, ...)
- v. Comparer le coût de soins par soignants (Docteur Jacques, Infirmier Dupond, ...)
- vi. Comparer le coût de soins par patients (âge, genre, métier, ...)
- vii. Comparer le coût de soins par responsables (Infirmière Dupuis, Docteur Jacques, ...)
- viii. Comparer le coût de soins par zones du corps (jambes, coeur, ...)

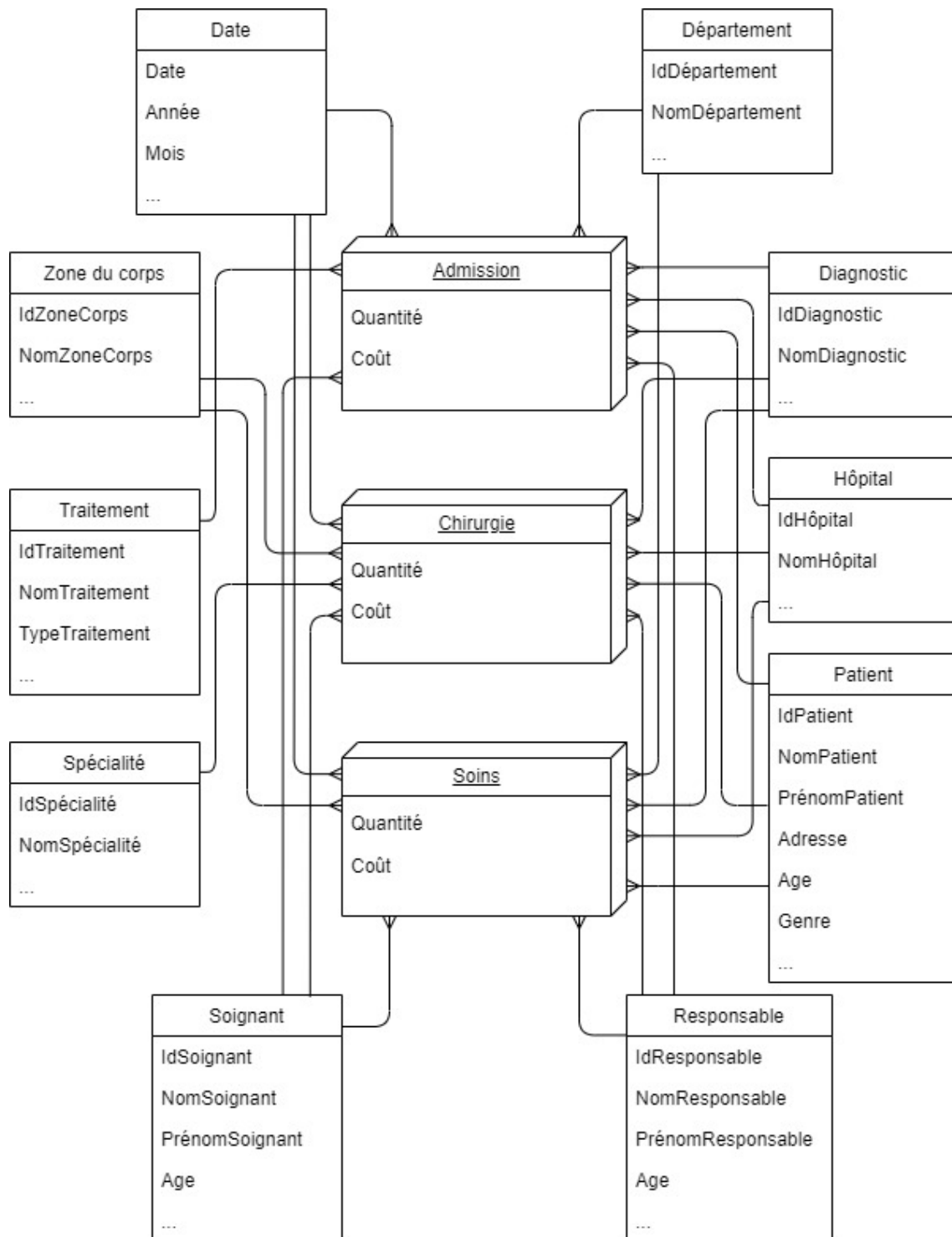
### Matrice de traduction des exigences

TABLE 7: Matrice de traduction des exigences du domaine médical

		Exigences					
		1a	1b	2a	2b	3a	3b
Faits et mesures	Admissions	✓	✓				
	Quantité	✓					
	Coût		✓				
	Chirurgie			✓	✓		
	Quantité			✓			
	Coût				✓		
	Soin					✓	✓
	Quantité					✓	
	Coût						✓
Dimensions	Date	✓	✓	✓	✓	✓	✓
	Département	✓	✓			✓	✓
	Diagnostic	✓	✓	✓	✓	✓	✓
	Hôpital	✓	✓	✓	✓	✓	✓
	Patient	✓	✓	✓	✓	✓	✓
	Responsable	✓	✓	✓	✓	✓	✓
	Soignant	✓	✓	✓	✓	✓	✓
	Spécialité			✓	✓		
	Traitement	✓	✓				
	Zone du corps			✓	✓	✓	✓

## Design pattern de domaine

FIGURE 9: Design pattern du domaine médical



### 3.4.7 Domaine de la culture

#### Liste des exigences

##### 1. Analyser les critiques :

###### (a) En terme de nombre :

- i. Analyser le nombre de critiques par acteurs (Will Smith, Jennifer Aniston, ...)
- ii. Analyser le nombre de critiques par auteurs (Pierre Dupuis, Jacques Dupond, ...)
- iii. Analyser le nombre de critiques par catégories (film, livre, ...)
- iv. Analyser le nombre de critiques par dates (année, mois, ...)
- v. Analyser le nombre de critiques par pays (Belgique, France, ...)
- vi. Analyser le nombre de critiques par produits critiqués (Harry Potter, )
- vii. Analyser le nombre de critiques par personnes réalisant les critiques (Marie Dupuis, Marc Dupond, ...)

###### (b) En terme de score :

- i. Comparer le score de critique par acteurs (Will Smith, Jennifer Aniston, ...)
- ii. Comparer le score de critiques par auteurs (Pierre Dupuis, Jacques Dupond, ...)
- iii. Comparer le score de critiques par catégories (film, livre, ...)
- iv. Comparer le score de critiques par dates (année, mois, ...)
- v. Comparer le score de critiques par pays (Belgique, France, ...)
- vi. Comparer le score de critiques par produits critiqués (Harry Potter, )
- vii. Comparer le score de critiques par personnes réalisant les critiques (Marie Dupuis, Marc Dupond, ...)

##### 2. Analyser les publications :

(a) En terme de nombre :

- i. Analyser le nombre de publications par auteurs (Pierre Dupuis, Jacques Dupond, ...)
- ii. Analyser le nombre de publications par catégories (articles de journaux, livre, ...)
- iii. Analyser le nombre de publications par dates (année, mois, ...)
- iv. Analyser le nombre de publications par éditeurs (Edition Lafont, ...)
- v. Analyser le nombre de publications par mots-clés (économie, Business Intelligence, ...)
- vi. Analyser le nombre de publications par pays (Belgique, France, ...)
- vii. Analyser le nombre de publications par publications produites (Larousse 2021, Le petit Prince, ...)

(b) En terme de longueur :

- i. Analyser la longueur des publications par auteurs (Pierre Dupuis, Jacques Dupond, ...)
- ii. Analyser la longueur des publications par catégories (articles de journaux, livres, ...)
- iii. Analyser la longueur des publications par dates (année, mois, ...)
- iv. Analyser la longueur des publications par éditeurs (Edition Lafont, ...)
- v. Analyser la longueur des publications par mots-clés (économie, Business Intelligence, ...)
- vi. Analyser la longueur des publications par pays (Belgique, France, ...)
- vii. Analyser la longueur des publications par publications produites (Larousse 2021, Le petit Prince, ...)

3. Analyser la réalisation de film :

(a) En terme de nombre :



- i. Analyser le nombre de films réalisés par acteurs (Will Smith, Jennifer Aniston, ...)
  - ii. Analyser le nombre de films réalisés par catégories (action, historique, ...)
  - iii. Analyser le nombre de films réalisés par dates (années, mois, ...)
  - iv. Analyser le nombre de films réalisés par pays (Belgique, France, ...)
  - v. Analyser le nombre de films réalisés par personnes de la production (scénariste, réalisateur, ...)
  - vi. Analyser le nombre de films réalisés par films produits (Titanic, Harry Potter, ...)
- (b) En terme de coût :
- i. Analyser le coût de films réalisés par acteurs (Will Smith, Jennifer Aniston, ...)
  - ii. Analyser le coût de films réalisés par catégories (action, historique, ...)
  - iii. Analyser le coût de films réalisés par dates (années, mois, ...)
  - iv. Analyser le coût de films réalisés par pays (Belgique, France, ...)
  - v. Analyser le coût de films réalisés par personnes de la production (scénariste, réalisateur, ...)
  - vi. Analyser le coût de films réalisés par films produits (Titanic, Harry Potter, ...)

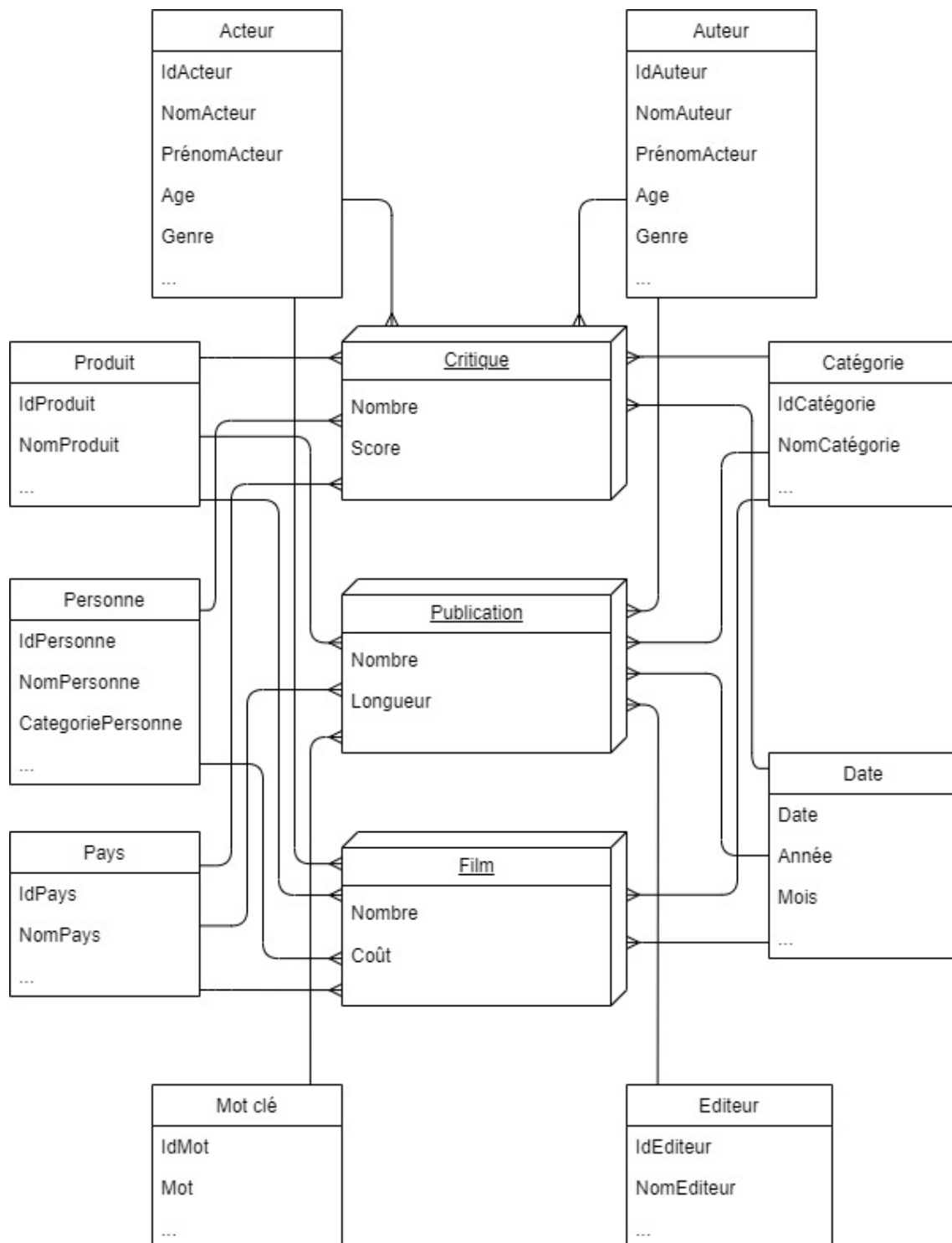
## Matrice de traduction des exigences

TABLE 8: Matrice de traduction des exigences du domaine de la culture

		Exigences					
		1a	1b	2a	2b	3a	3b
Faits et mesures	Critique	✓	✓				
	Nombre	✓					
	Score		✓				
	Publication			✓	✓		
	Nombre			✓			
	Longueur				✓		
	Film					✓	✓
	Nombre					✓	
	Coût						✓
Dimensions	Acteur	✓	✓			✓	✓
	Auteur	✓	✓	✓	✓		
	Catégorie	✓	✓	✓	✓	✓	✓
	Date	✓	✓	✓	✓	✓	✓
	Editeur			✓	✓		
	Mot clé			✓	✓		
	Pays	✓	✓	✓	✓	✓	✓
	Personne	✓	✓			✓	✓
	Produit	✓	✓	✓	✓	✓	✓

## Design pattern de domaine

FIGURE 10: Design pattern du domaine de la culture



### 3.4.8 Domaine de la géologie

#### Liste des exigences

##### 1. Analyser l'érosion :

###### (a) En terme de nombre :

- i. Comparer le nombre d'érosions par cours d'eau (Sambre, Eau d'Heure, ...)
- ii. Comparer le nombre d'érosions par dates (année, mois, ...)
- iii. Comparer le nombre d'érosions par localisations (pays, région, ...)
- iv. Comparer le nombre d'érosions par météo (précipitation, température, ...)
- v. Comparer le nombre d'érosions par risques (pente, distance, type de berge, ...)
- vi. Comparer le nombre d'érosions par utilisation des endroits (parc, forêt, ...)

###### (b) En terme de taille :

- i. Comparer la taille des érosions par cours d'eau (Sambre, Eau d'Heure, ...)
- ii. Comparer la taille des érosions par dates (année, mois, ...)
- iii. Comparer la taille des érosions par localisations (pays, région, ...)
- iv. Comparer la taille des érosions par météo (précipitation, température, ...)
- v. Comparer la taille des érosions par risques (pente, distance, type de berge, ...)
- vi. Comparer la taille des érosions par utilisations des endroits (parc, forêt, ...)

##### 2. Analyser la pollution de l'eau :

###### (a) En terme de nombre :

- i. Comparer le nombre de pollution par cours d'eau (Sambre, Eau d'Heure, ...)
- ii. Comparer le nombre de pollution par dates (années, mois, ...)
- iii. Comparer le nombre de pollution par localisations (pays, région, province, ...)

- iv. Comparer la nombre de pollution par produits polluants (pesticides, hydrocarbures, ...)
- (b) En terme d'intensité :
- i. Comparer l'intensité de pollution par cours d'eau (Sambre, Eau d'Heure, ...)
  - ii. Comparer l'intensité de pollution par dates (années, mois, ...)
  - iii. Comparer l'intensité de pollution par localisations (pays, région, province, ...)
  - iv. Comparer l'intensité de pollution par produits polluants (pesticides, hydrocarbures, ...)
3. Analyser la production agricole :
- (a) En fonction de la quantité récoltée :
- i. Comparer la quantité récoltée par productions agricoles par dates (année, mois, ...)
  - ii. Comparer la quantité récoltée par productions agricoles par localisations (région, province, ...)
  - iii. Comparer la quantité récoltée par productions agricoles par météos (précipitations, température, ...)
  - iv. Comparer la quantité récoltée par productions agricoles par produits cultivés (betteraves, maïs, ...)
  - v. Comparer la quantité récoltée par productions agricoles par types de cultures (traditionnel, bio, ...)
- (b) En fonction de la superficie cultivée :
- i. Comparer la superficie cultivée par productions agricoles par dates (année, mois, ...)
  - ii. Comparer la superficie cultivée par productions agricoles par localisations (région, province, ...)

- iii. Comparer la superficie cultivée par productions agricoles par météo (précipitation, température, ...)
- iv. Comparer la superficie cultivée par productions agricoles par produits cultivés (betteraves, maïs, ...)
- v. Comparer la superficie cultivée par productions agricoles par types de cultures (traditionnel, bio, ...)

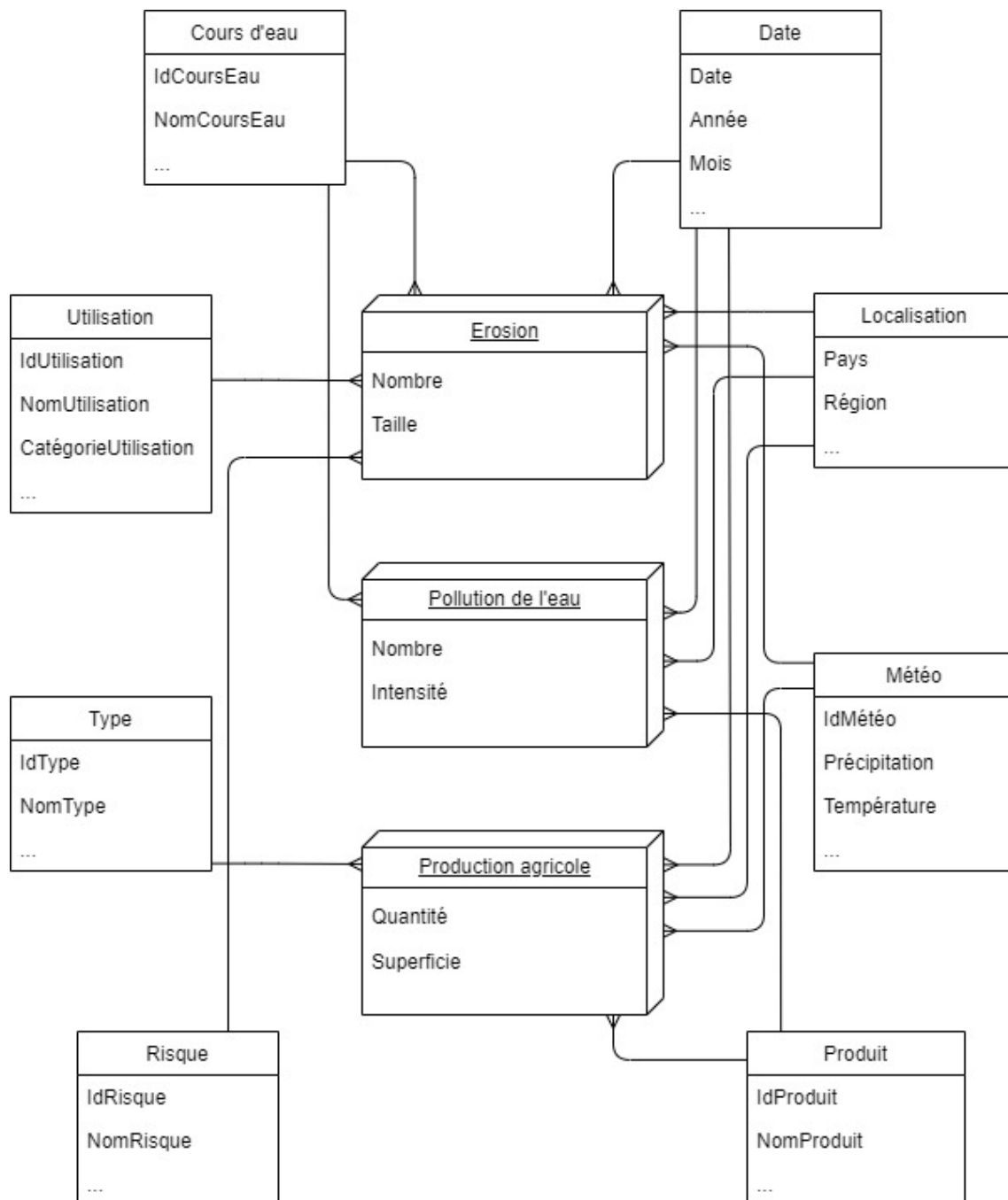
### Matrice de traduction des exigences

TABLE 9: Matrice de traduction des exigences du domaine géologie

		Exigences					
		1a	1b	2a	2b	3a	3b
Faits et mesures	Erosion	✓	✓				
	Nombre	✓					
	Taille		✓				
	Pollution de l'eau			✓	✓		
	Nombre			✓			
	Intensité				✓		
	Production agricole					✓	✓
	Quantité					✓	
	Superficie						✓
Dimensions	Cours d'eau	✓	✓	✓	✓		
	Date	✓	✓	✓	✓	✓	✓
	Localisation	✓	✓	✓	✓	✓	✓
	Météo	✓	✓			✓	✓
	Produit			✓	✓	✓	✓
	Risque	✓	✓				
	Type					✓	✓
	Utilisation	✓	✓				

## Design pattern de domaine

FIGURE 11: Design pattern du domaine géologie



## 4 Résultats

Le concept de design pattern a donc été adapté au cas de la conception de Data warehouses par le biais des domaines design patterns accompagnés de listes d'exigences et de matrices de traduction. Les résultats de cette application sont abordés dans ce chapitre. Le premier point concerne l'application de cette adaptation. Par la suite, les limites de l'application sont abordées. Le dernier point consiste en une discussion sur les pistes d'améliorations de cette utilisation du concept.

### 4.1 Application

Dans l'optique de faciliter la conception de nouveaux Data Warehouses, l'objectif de ce travail est d'identifier et de représenter les caractéristiques communes aux conceptions de Data Warehouses via les design patterns. L'utilisation de la méthode se fait donc lors d'une nouvelle conception.

Préalablement à l'utilisation de cette méthode, il faut déterminer le sujet central et les besoins de l'analyse. La première étape de la méthode consiste à identifier le domaine d'étude le plus adéquat en fonction du sujet d'analyse. La deuxième étape est d'identifier, dans la liste des exigences pré-établies du domaine, les exigences souhaitées. La troisième étape consiste à identifier, dans la matrice de traduction, les dimensions et mesures liées aux exigences souhaitées. La quatrième étape est d'identifier le fait, les mesures et les dimensions dans le design pattern du domaine. La cinquième étape est de créer le schéma conceptuel du Data Warehouse sur base des éléments identifiés précédemment.

Pour illustrer l'utilisation de cette méthode, une démonstration à l'aide d'un exemple est de mise. Imaginons que l'hôpital Sainte-Elisabeth de Namur souhaite en savoir plus à propos de ses opérations. Elle souhaite les analyser et les comparer en terme de nombre et de coût selon les dates, les médecins, les patients, les spécialités et les zones du corps. Premièrement, on peut établir que ce sujet d'étude est relatif au domaine médical. Deuxièmement,



xièmement, en reprenant la liste des exigences pré-établies pour ce domaine, on constate que cet exemple repose sur le deuxième type d'exigences (2) concernant les chirurgies selon les mesures de nombre (a) et de coût (b). Plus précisément, l'hôpital souhaite analyser les points concernant les dates (i), les médecins (iv), les patients (v), les spécialités (vii) et les zones du corps (viii). Troisièmement, comme montré à la figure 10, on identifie dans la matrice de traduction le fait, les mesures et les dimensions que les exigences souhaitées requièrent. Quatrièmement, comme montré à la figure 12, le fait et les dimensions identifiées à l'étape précédente sont identifiés dans le design pattern du domaine. Cinquièmement, en reprenant les éléments identifiés dans le domaine design pattern à l'étape précédente, le schéma conceptuel du Data Warehouse nécessaire à l'analyse souhaitée par l'hôpital est représenté à la figure 13.

TABLE 10: Exemple d'utilisation de la matrice de traduction des exigences du domaine médical

		Exigences					
		1a	1b	2a	2b	3a	3b
Faits et mesures	Admissions	✓	✓				
	Quantité	✓					
	Coût		✓				
	Chirurgie			✓	✓		
	Quantité			✓			
	Coût				✓		
	Soin					✓	✓
	Quantité					✓	
	Coût						✓
Dimensions	Date	✓	✓	✓	✓	✓	✓
	Département	✓	✓			✓	✓
	Diagnostic	✓	✓	✓	✓	✓	✓
	Hôpital	✓	✓	✓	✓	✓	✓
	Patient	✓	✓	✓	✓	✓	✓
	Responsable	✓	✓	✓	✓	✓	✓
	Soignant	✓	✓	✓	✓	✓	✓
	Spécialité			✓	✓		
	Traitement	✓	✓				
	Zone du corps			✓	✓	✓	✓

FIGURE 12: Exemple d'utilisation du design pattern du domaine médical

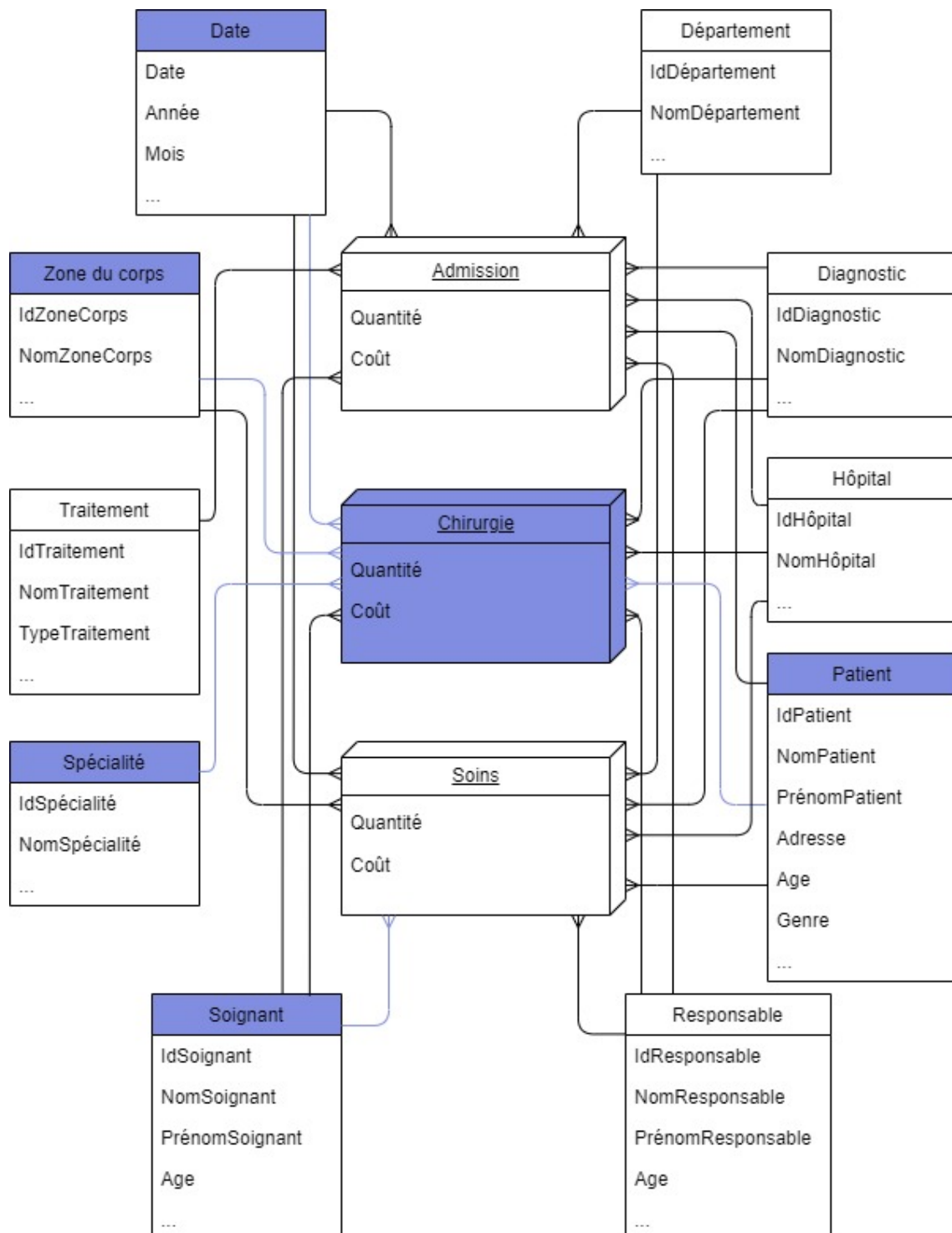
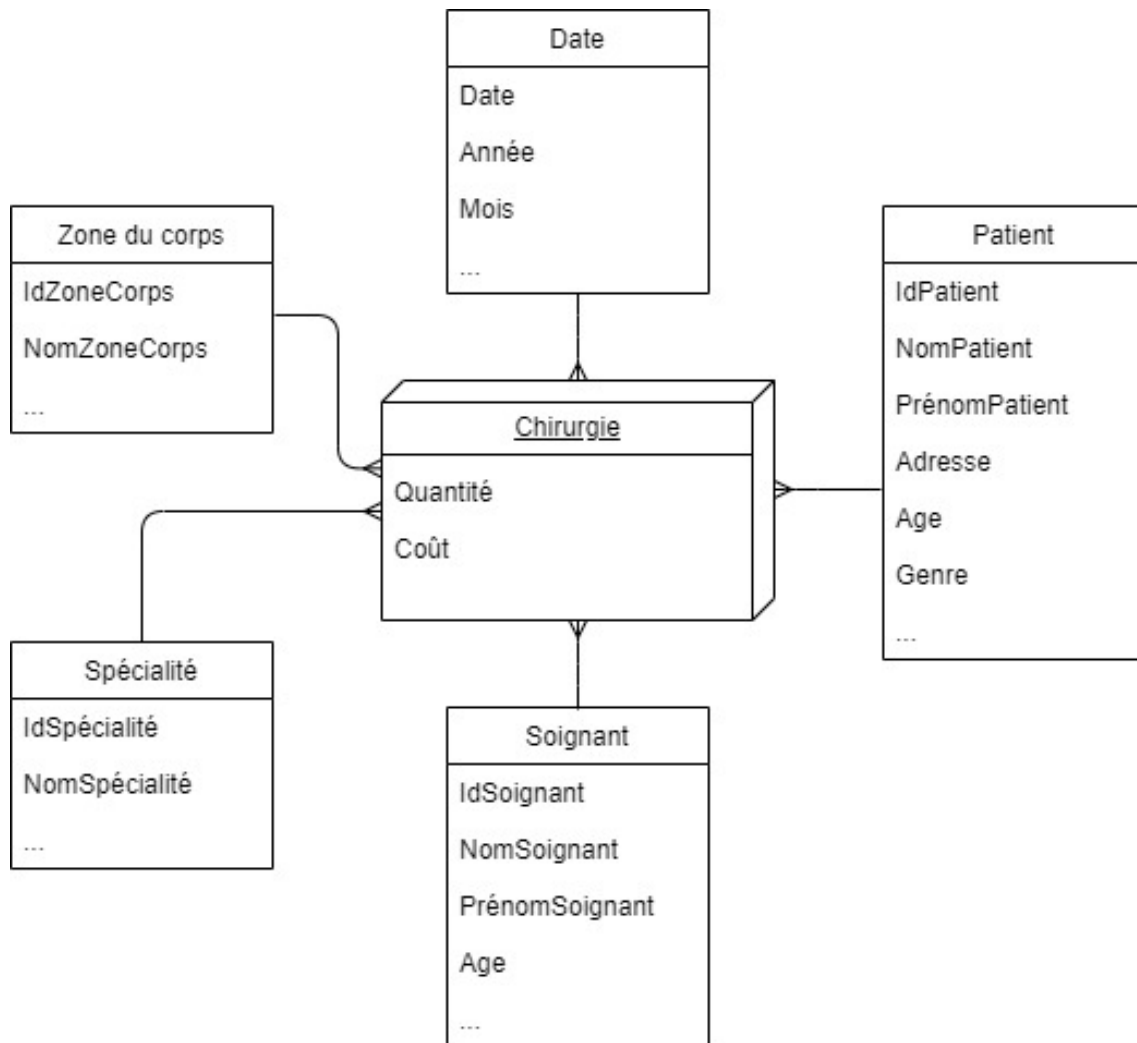


FIGURE 13: Exemple de schéma conceptuel de Data Warehouse



## 4.2 Limites

L'application de ces domaines design patterns a cependant certaines limites. Tout d'abord, ce mémoire réalise la première application du concept de design patterns au cas de la conception de Data Warehouses. Comme on peut le constater avec le cas des bases de données classiques, un certain temps est nécessaire pour que tout le monde s'accorde sur les design patterns d'un domaine. Au fur et à mesure des utilisations de la méthode ainsi que des futures recherches sur ce sujet, les design patterns identifiés et représentés

dans ce mémoire seront amenés à évoluer.

Ensuite, le temps nécessaire à l'application d'un design pattern peut être sous-estimé. En effet, c'est une erreur de penser que la conception à l'aide de design patterns prend peu de temps. Tout d'abord, le développeur doit prendre du temps pour comprendre la méthode ainsi que ses trois composantes. Par la suite, il doit également prendre le temps de décider comment l'appliquer et l'adapter à son cas d'étude et ses besoins d'analyses.

Par la suite, comme déjà mentionné précédemment, l'approche idéale de baser les design patterns sur des éléments transversaux à toutes les conceptions tout comme en ingénierie logicielle a été difficilement accessible. L'absence de tels éléments est probablement dû à la collecte de données. En effet, on peut constater que les cas sont fort théoriques et pas très complexes. Cela peut donc expliquer le manque d'éléments transversaux.

L'identification des exigences par domaines présente également certaines limites. En effet, comme dit précédemment, les listes d'exigences sont non exhaustives car elles ont été élaborées à l'aide des informations disponibles dans les différentes sources des schémas collectés ainsi qu'avec mes connaissances issues de ma formation en ingénieur de gestion. Il semble que la qualité et complétude de ces listes pourraient être améliorées.

Dernièrement, ni les exigences ni les domaines design patterns ne font mention de hiérarchies. Celles-ci sont en effet absentes de la méthode. Elles ont à chaque fois été normalisées, similairement aux représentations en étoile, dans la dimension la plus haute de la hiérarchie. C'est un point manquant de ce travail.

### **4.3 Discussion**

Pour aller plus loin et compléter ce travail, plusieurs éléments peuvent être envisagés. Tout d'abord, les design patterns identifiés et représentés pourraient être complétés. En effet, les hiérarchies les plus courantes pourraient être identifiées. Une idée pourrait être d'offrir deux alternatives. Il y aurait alors un design pattern avec hiérarchies et un

sans hiérarchies. Il y aurait de cette manière moins de travail pour le développeur. Une autre idée pourrait être de garder les domaines design patterns identiques mais d'ajouter une colonne dans les matrices pour indiquer les dimensions pouvant comprendre des hiérarchies.

Le problème de complétude des listes d'exigences est une des limites qui pourrait le plus facilement être résolue. De plus, il semble que cela améliorerait l'applicabilité de la méthode. Pour se faire, l'idée privilégiée est de réaliser des interviews avec des experts par domaines. Ces experts semblent être plus à même d'identifier des exigences plus complètes et pointues.

En ce qui concerne la problématique des éléments transversaux dans les conceptions, elle pourrait être potentiellement résolue avec une collecte alternative de données. En effet, il semble qu'une collecte de données en entreprises donnerait des cas plus complexes. L'analyse de tels cas permettrait peut-être d'identifier des éléments transversaux et d'appliquer le concept de design pattern de manière similaire à l'application réalisée dans le domaine de l'ingénierie logicielle.

Dernièrement, pour aller plus loin dans l'aide à la conception de Data Warehouses, une possibilité peut être d'accompagner les design patterns d'un recueil de schémas. Il est évident que des schémas sont disponibles à travers des livres, des articles ou sur internet mais aucun inventaire n'existe. Celui-ci permettrait aux développeurs de disposer à un endroit unique de nombreux exemples et alternatives. Ils pourraient donc plus facilement s'inspirer de conceptions existantes. Cet inventaire compléterait bien les domaines design patterns qui peuvent sembler abstraits.

## 5 Conclusion

L'importance du domaine de la Business Intelligence n'est plus à prouver. Pour parvenir à une telle solution, nul ne peut se passer de Data Warehouses. Des méthodes de conception formelles permettent de les développer mais elles présentent des difficultés. Dans d'autres domaines, des méthodes de résolution éprouvées ont fait leur preuve. Pourquoi ne pas les adapter aux cas de la conception de Data Warehouses ?

C'est dans le domaine de l'ingénierie logicielle ainsi que dans celui des bases de données qu'il semble le plus pertinent de faire des recherches liées aux méthodes de résolution de problèmes. Avec les nombreux avantages qu'il offre, le concept de design patterns est utilisé dans ces deux domaines. De plus, ce concept est facilement modulable selon les domaines. L'adaptation de ce concept au cas de la conception de Data Warehouses semble donc judicieuse.

Pour parvenir à une telle adaptation, la méthodologie a été la suivante. Tout d'abord, une collecte de schémas de conception de Data Warehouses a été réalisée. Trois types de sources ont été utilisées pour assurer la qualité et la variété des schémas. Ceux-ci peuvent être considérés comme des solutions éprouvées de conception et ont donc servi de base à l'identification de design patterns. Par la suite, l'analyse de ces schémas a permis d'identifier huit domaines métier dans lesquels presque tous les schémas se retrouvent. Sur base de cette analyse, pour le cas de la conception de Data Warehouses, le concept a été adapté en domaine design pattern. Ces domaines design patterns ont été identifiés, représentés et accompagnés d'une liste d'exigences et d'une matrice de traduction. Pour clarifier au mieux l'utilisation de ces domaines design patterns, un exemple d'application au domaine médical a été réalisé.

Cette adaptation a néanmoins certaines limites telles que le caractère non exhaustif de la liste d'exigences ou encore l'absence de hiérarchies. Heureusement, des solutions potentielles peuvent être identifiées pour la majorité des limites et ne demandent qu'à être développées.

## Annexe

En annexe se trouvent d’abord les tables reprenant tous les schémas analysés dans ce travail, en fonction des types de sources. Pour chacun, il est mentionné le titre, l’auteur, la référence et la page à laquelle le trouver. A la suite de ces tables, sont repris tous les schémas recueillis.

TABLE 11: Schémas extraits des livres de références

Schéma	Titre	Auteurs	Référence	Page
1	Data Warehouses Systems	Alejandro Vaisman	[1]	78
2		Esteban Zimanyi		79
3				80
4				81
5				82
6				83
7				84
8	The Data Warehouse Toolkit (Third Edition)	Ralph Kimball	[2]	85
9		Margy Ross		85
10				86
11				86
12				87
13				88
14				88
15				89
16				89
17				90
18				90
19				91
20				92

21				93
22				93
23				94
24				94
25	Data Warehouse Design :	Matteo Golfarelli	[3]	95
26	Modern principles and	Stefano Rizzi		95
27	methodologies			96
28				97
29				97
30				98
31				98
32				99
33				99
34				100
35				100

TABLE 12: Schémas extraits d'articles scientifiques

Schéma	Titre	Auteurs	Référence	Page
36	A Markup Language for Building a Data Warehouse Warehouse or Educational Peer-Assessment Research	Yang Song Ferry Pramudianto Edward F Gehringer	[28]	101
37	A service oriented approach to Business Intelligence in Telecoms industry	Tanko Ishaya Musiliudeen Folarin	[29]	102
38	Academic Data Warehouse	Ezlo Lefons	[30]	103



39	Design Using a Hybrid Methodology	Filippo Tangorra Tria Francesco Di		104
40	Agile Development for Service Business Intelligence Solutions	Marina Mircea Anca Iona Andreescu	[31]	105
41	Analyse du potentiel d'une application SOLAP pour une gestion efficace de l'érosion des berges en Gaspésie Iles-de-la-Madeleine	Sonia Rivest Yvan Bédard Michel Michaud Rosemarie McHugh Francis Bilodeau	[32]	106
42	Analyse multi-dimensionnelle de documents via des dimensions OLAP	Franck Ravat Olivier Teste Ronan Tournier	[33]	107
43	Analysis of Student Data for Retention Using Data Mining Techniques	Brandon Sherrill William Eberle Douglas A. Talbert	[34]	108
44	Automating Data Warehouse Conceptual Schema Design and Evaluation	Cassandra Phipps Karen C. Davis	[35]	109
45	Business Intelligence approach to supporting strategy-making of ISP service management	Sheng-Tun Li Li-Yen Shue Shu-Fen Lee	[36]	110
46	Business Intelligence for product defect	A S Girsang S M Isa	[37]	111

	analysis	A L Haris Arwan K Mandagie L R Ariana V Ardinda		
47	Using geospatial business intelligence paradigm to design a multidimensional conceptual model for efficient coastal erosion risk assessment	Amaneh Jadidi Mir Abolfazl Mostafavi Yvan Bédard Bernard Long Eve Grenier	[38]	112
48	Towards of a Business	Danilo Coelho	[39]	113
49	Intelligence Platform to	João Miranda		113
50	Portuguese Misericórdias	Filipe Portela José Machado Manuel Filipe Santos António Abelha		114
51	Towards Exploratory OLAP over Linked Open Data : A Case Study	Dilshod Ibragimov Katja Hose Torben Bach Pedersen Esteban Zimanyi	[40]	114
52	starER : A Conceptual Model for Data Warehouse Design	Nectaria Tryfona Frank Busborg Jens G. Borch Christiansen	[41]	115
53	QBX : A CASE Tool for Data Mart Design	Antonino Battaglia Matteo Golfarelli Stefano Rizzi	[42]	116

54	Configurable Web	Andrea Delgado	[43]	117
55	Warehouse construction through BPM Systems	Adriana Marotta		118
56	The concept of document warehousing for multi-dimensional modeling of textual-based business intelligence	Frank S.C. Tseng Annie Y.H. Chou	[44]	119 120 120
59	Metrics for data warehouse conceptual models understandability	Manuel Serrano Juan Trujillo Coral Calero Mario Piattini	[45]	121
60	Metadata Management in a Multiversion Data Warehouse	Robert Wrembel Bartosz Bebel	[46]	122
61	Implementation of Time Dimension in Star Schema Design : A Case Study of Birth Registration Data	Pushpal Desai	[47]	123
62	From conceptual models to schemata : An object-process-based data warehouse construction method	Dov Dori Roman Feldman Arnon Sturm	[48]	124
63	Framework for Designing a Healthcare Outcome Data Warehouse	Bambang Parmanto Matthew Scotch Sjarif Ahmad	[49]	125
64	Evaluating a Healthcare	Osama E. Sheta	[50]	126

	Data Warehouse For Cancer Diseases	Ahmed Nour Eldeen		
65	Designing Data Warehouses with OO Conceptual Models	Juan Trujillo Manuel Palomar Jaime Gomez Il-Yeol Song	[12]	127
66 67	Design of a data warehouse model for a university decision support system	Ibrahim Inuwa N. D. Oye	[51]	127 128
68	Data Warehouse performance - selected techniques and data structures	Robert Wrembel	[52]	128
69	Data warehouse design from social media for opinion analysis : the case of Facebook and Twitter	Imen Moalla Ahlem Nabli Lotfi Bouzguenda Mohamed Hammami	[53]	129
70	Conceptual Data Warehouse Design	Bodo Hüsemann Jens Lechtenbörger Gottfried Vossen	[11]	130
71	Creating business intelligence from course management systems	Liezl van Dyk Pieter Conradie	[54]	130
72 73	Hybrid methodology for data warehouse conceptual design by UML schemas	Francesco Di Tria Ezio Tria Filippo Tangorra	[55]	131 131

74	Querying data warehouses efficiently using the Bitmap Join Index OLAP Tool	Anderson Chaves Carniel Thiago Luís Lopes Siqueira	[56]	132
75	Data warehouses : next challenges	Alejandro Vaisman Esteban Zimanyi	[57]	133 134
77	Logical representation of a conceptual model for Spatial Data Warehouse	Elzbieta Malinowski Esteban Zimanyi	[58]	135
78	Conceptual modeling for data warehouse and OLAP applications	Elzbieta Malinowski Esteban Zimanyi	[59]	136 137
80	Goal-Oriented requirement analysis for data warehouse design	Paolo Giorgini	[60]	138
81	A Data Warehouse/OLAP Framework for Web Usage Mining and Business Intelligence Reporting	Xiaohua Hu Nicke Cerccone	[61]	139
82	OMC-IDS at the cross-roads of OLAP mining and intrusion detection	Hanen Brahmi Imen Brahmi Sadok Ben Yahia	[62]	140
83	A proposed model for data warehouse ETL processes	S. H. A. El-Sappagh A. M. A. Hendawi A. H. El Bastawissy	[63]	140
84	A UML-based data warehouse design method	Nicolas Prat Jacky Akoka Isabelle Comyn-Wattiau	[64]	141

85	A Methodological Framework for Data Warehouse Design	Matteo Golfarelli Stefano Rizzi	[6]	142
----	---	------------------------------------	-----	-----

TABLE 13: Schémas extraits de sites internet

Schéma	Titre	Auteurs	Référence	Page
86	Understand star schema and the importance for Power BI	Peter Myers Kate Follis Kent Sharkey Maggie Sparkman Michael Blythe David Coulter Adam Saxton	[65]	143
87	SQL Server Data	Robert Sheldon	[66]	144
88	Warehouse Cribsheet			145
89	Data Warehouse : basics of relational vs star schema data modeling	Daryl Ung	[67]	146
90	Comprehensive guide to the Data Warehouse	Nicolas Janeway Bills	[68]	147
91	Modeling your dimen-	Iqbal Ahmed Alvi	[69]	148
92	sional Data Warehouse			149
93	Conception d'un entre-	Yazid Grim	[70]	150
94	pôt de données (Data Warehouse)	Fleur-Anne Blain		151
95	Quels sont les buts recherchés et les enjeux d'un DataWeb ?	Sylvain Bourdette	[71]	152

96	Exam Prep 70-463 : Tables and Schemas	Colleen M. Morrow	[72]	153
97 98	Data preparation : first and foremost important task in Power BI	Reza Rad	[73]	154 155
99 100	Star and snowflake schema in data warehousing	Enrico Besenyei	[74]	156 157

Schéma 1

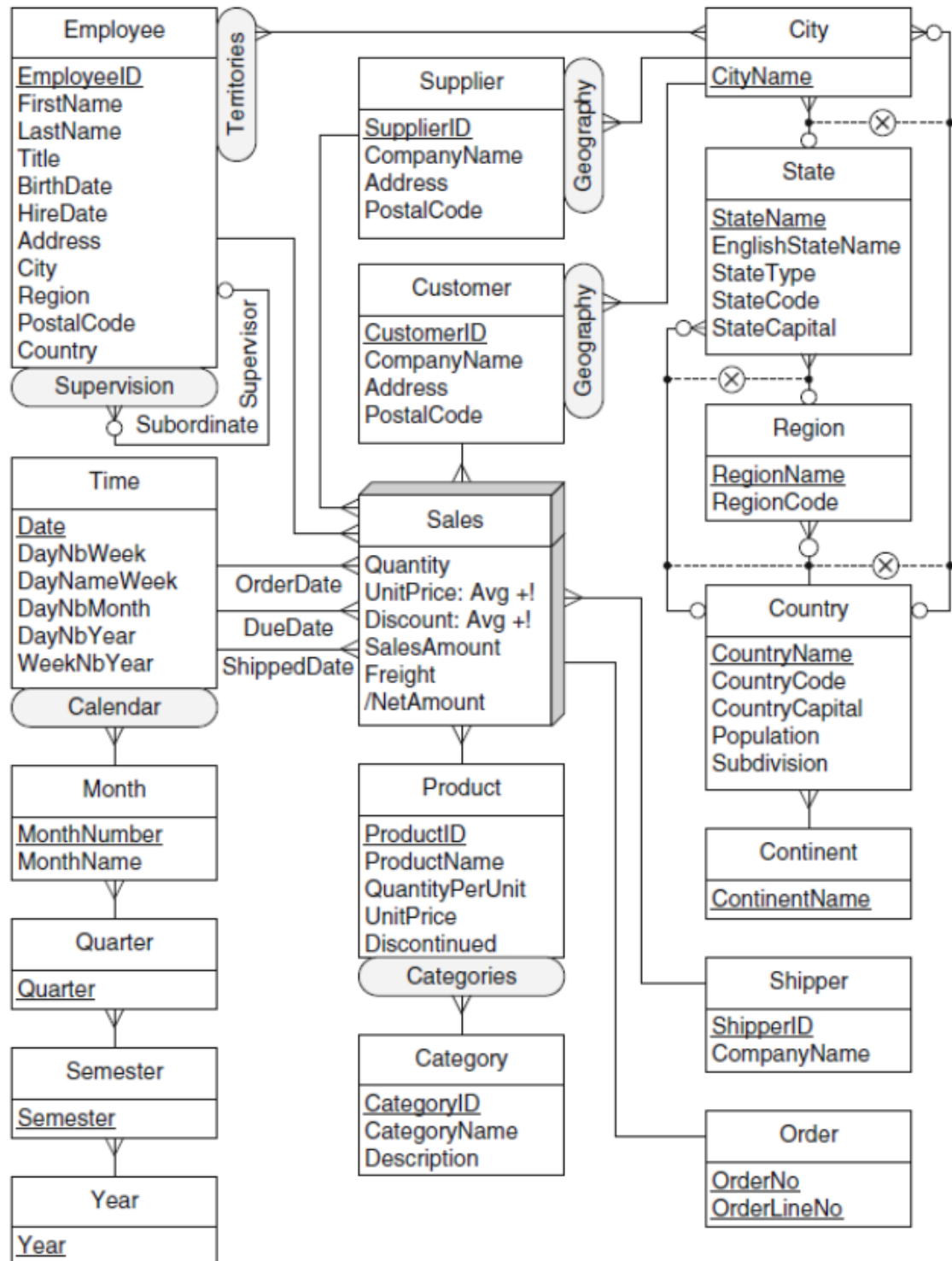




Schéma 2

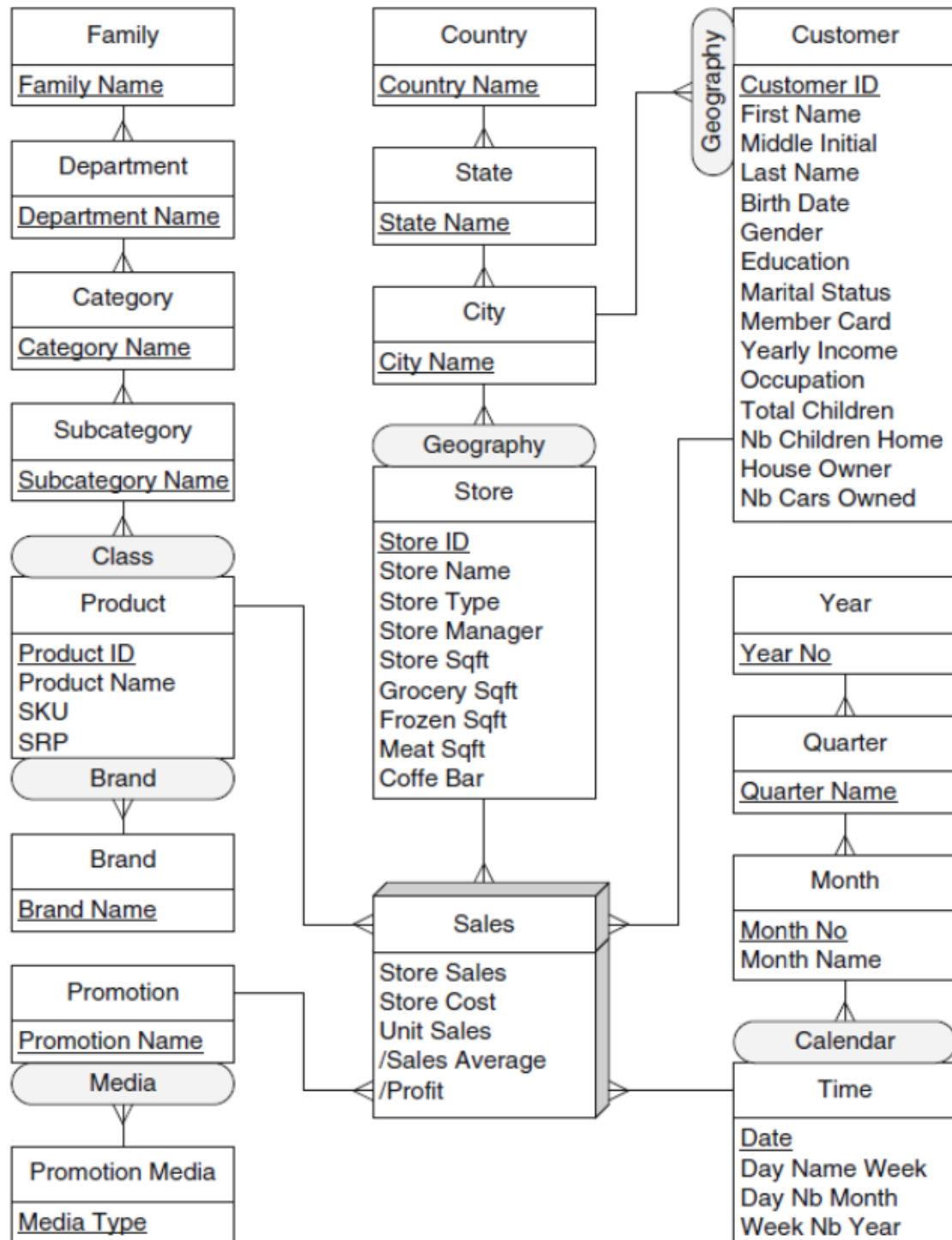


Schéma 3

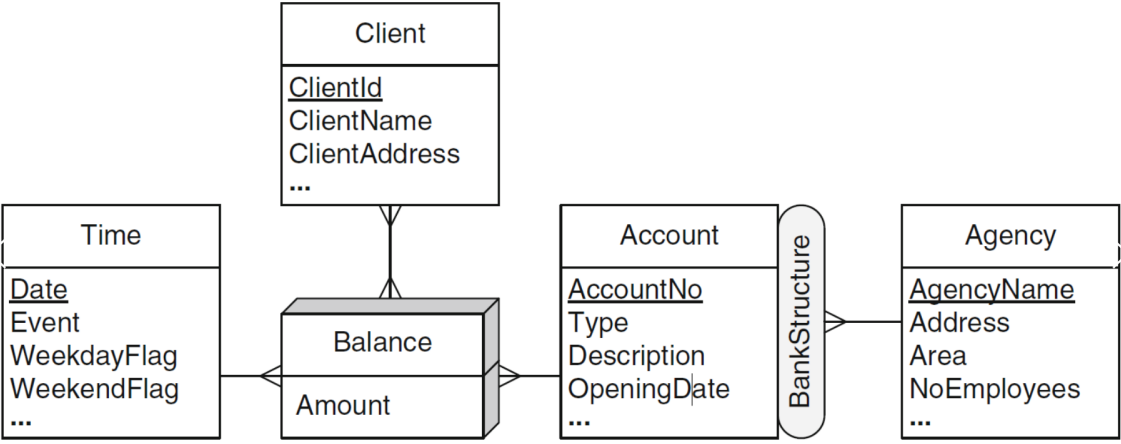


Schéma 4

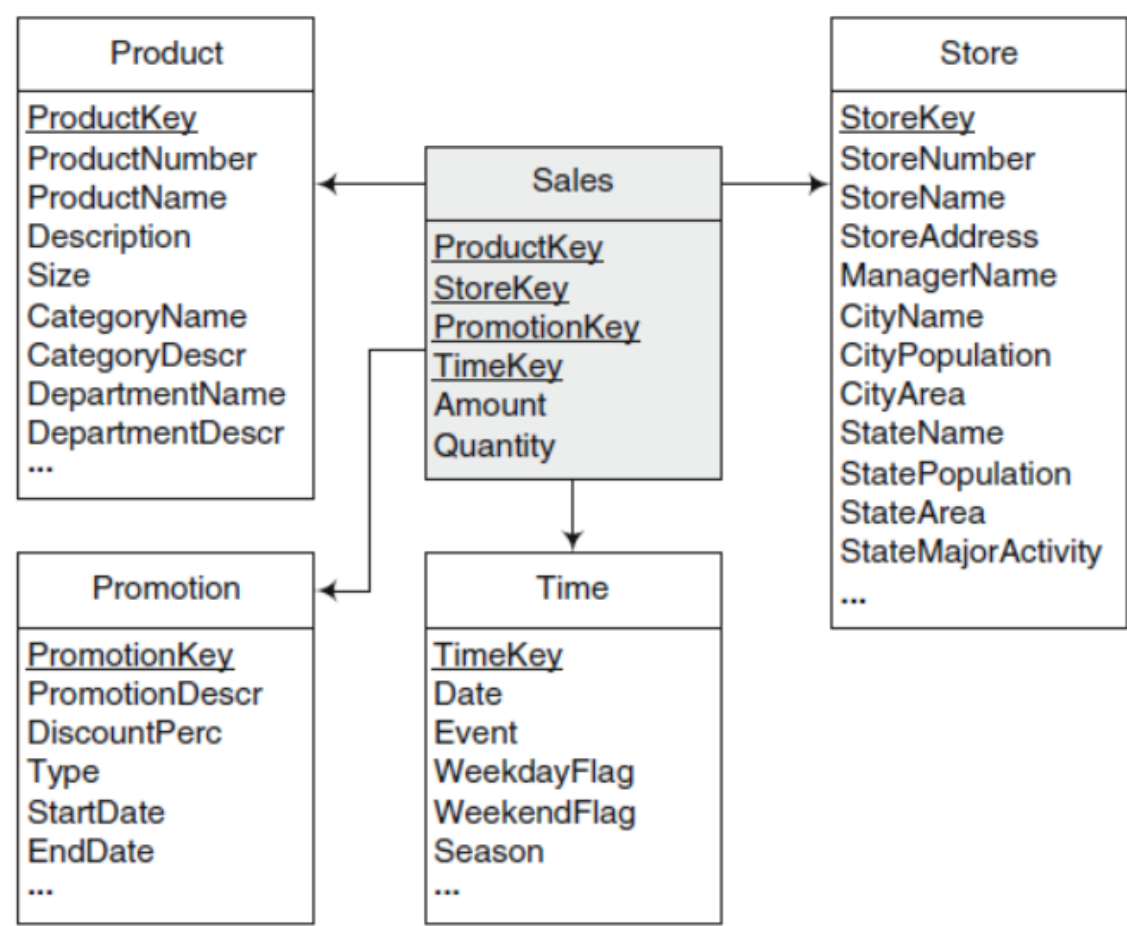


Schéma 5

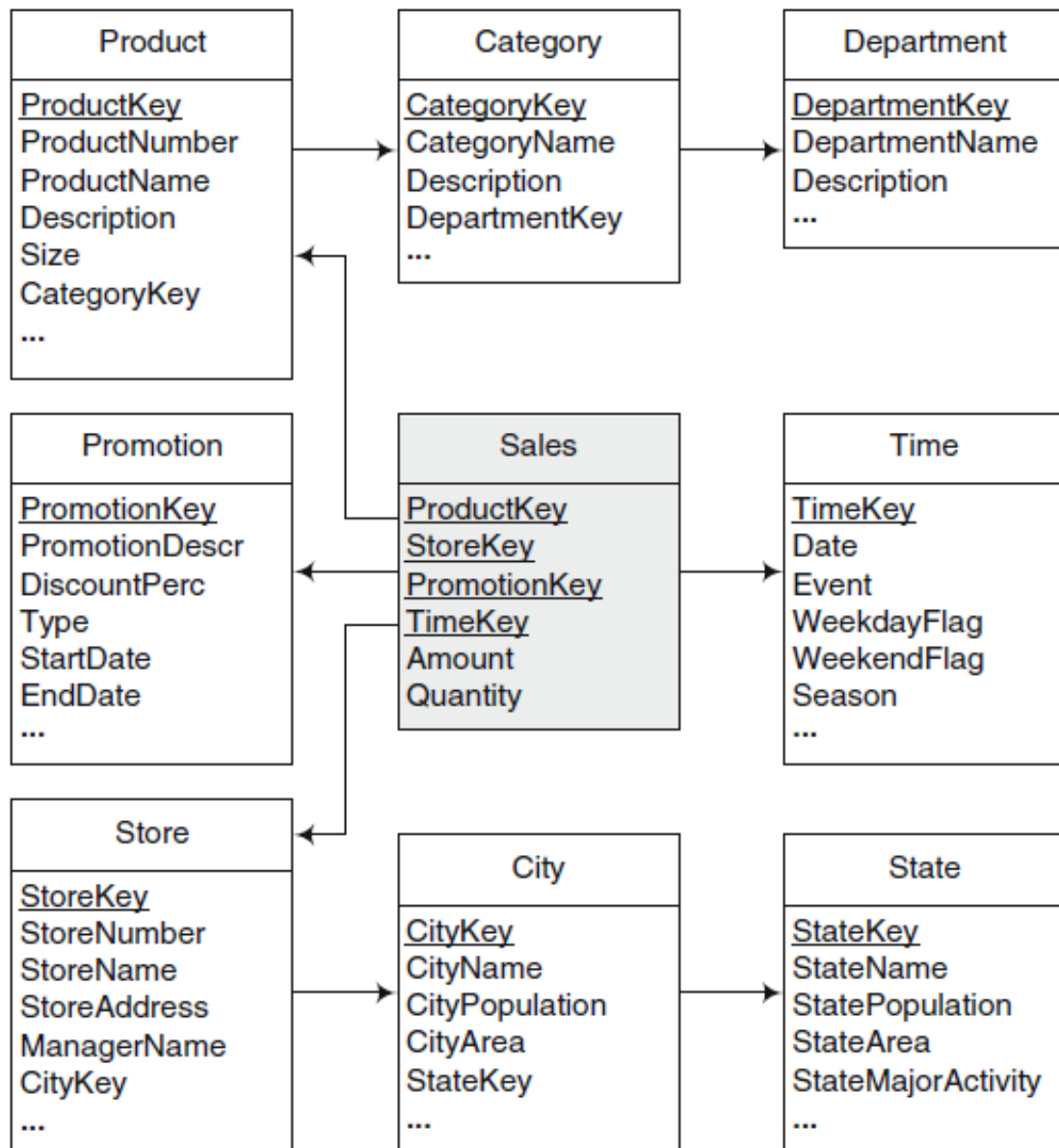


Schéma 6

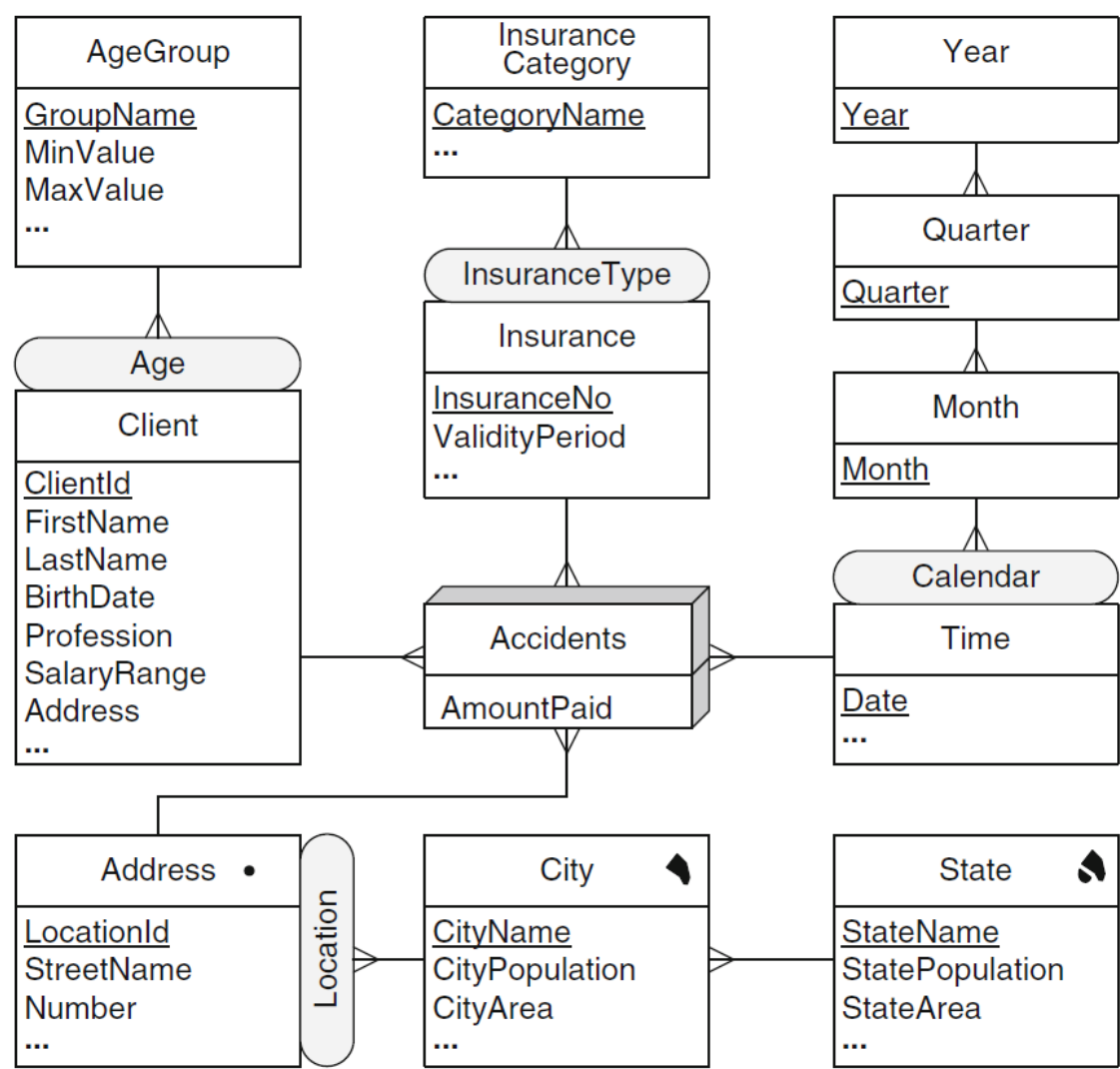


Schéma 7

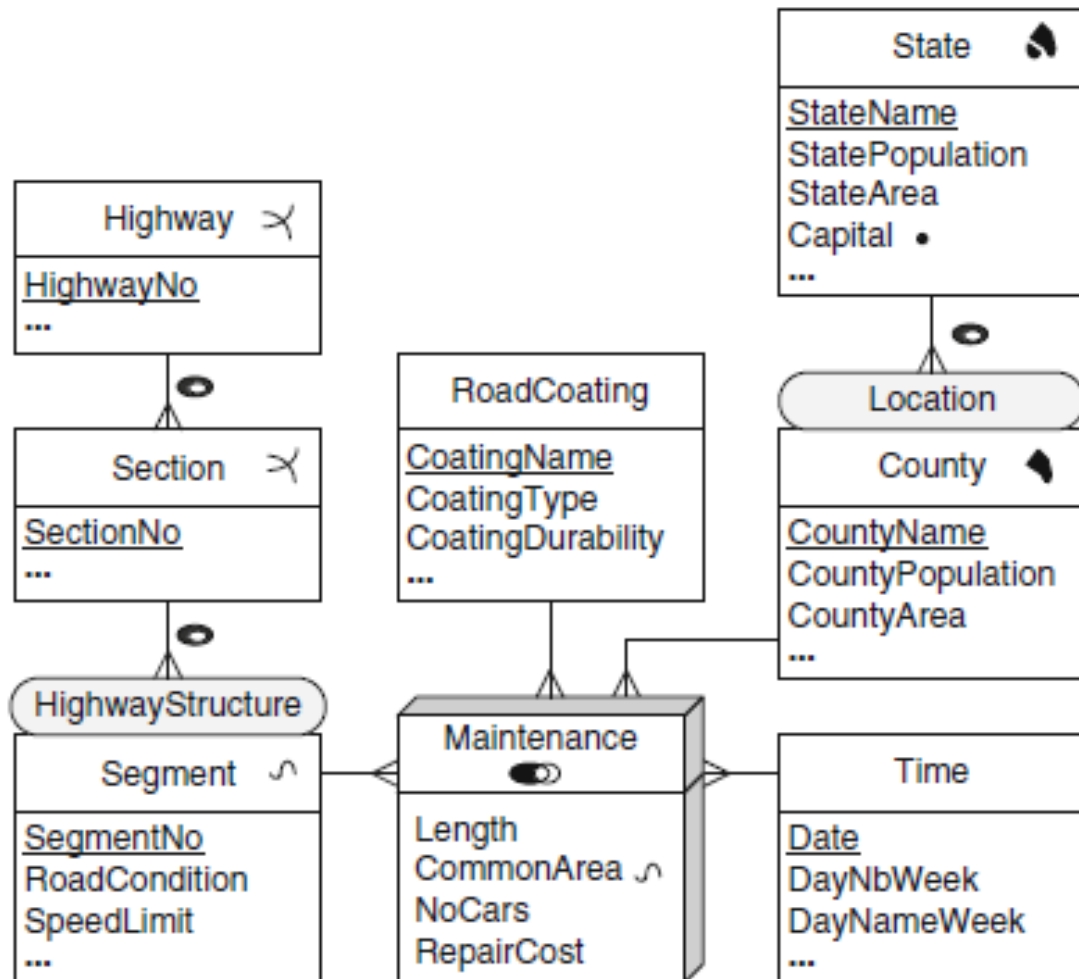


Schéma 8

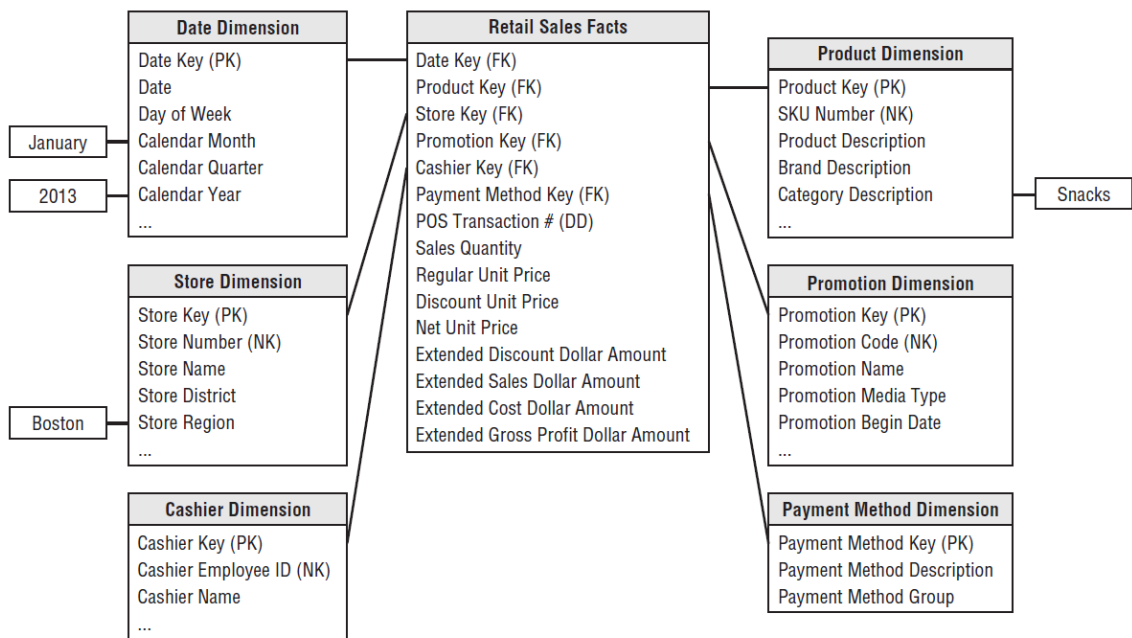


Schéma 9



Schéma 10

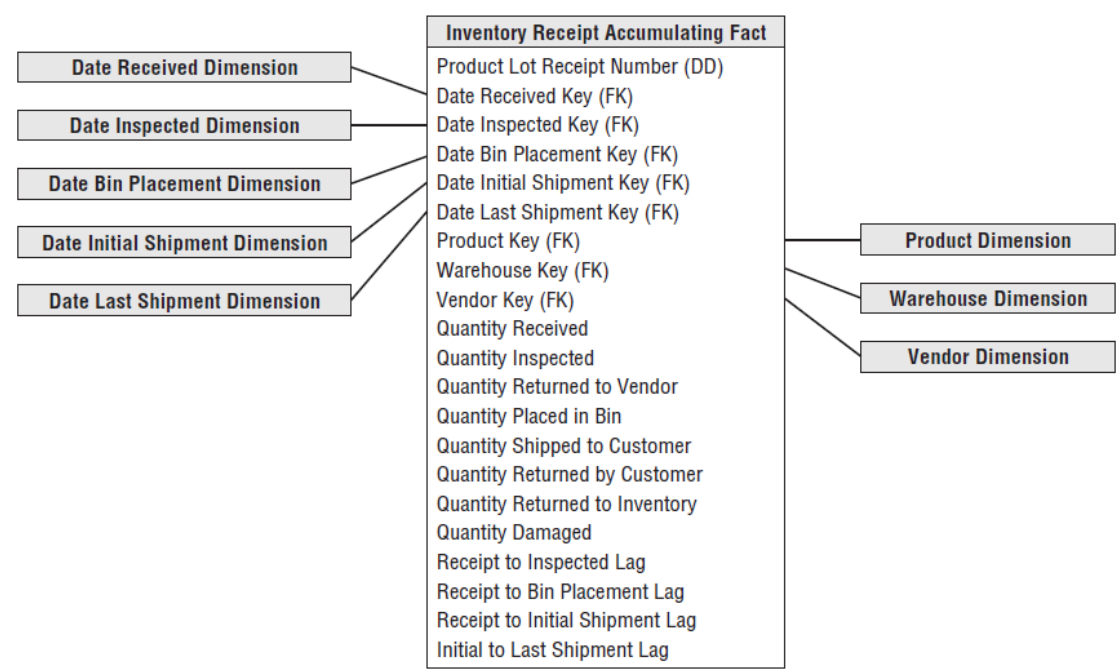


Schéma 11

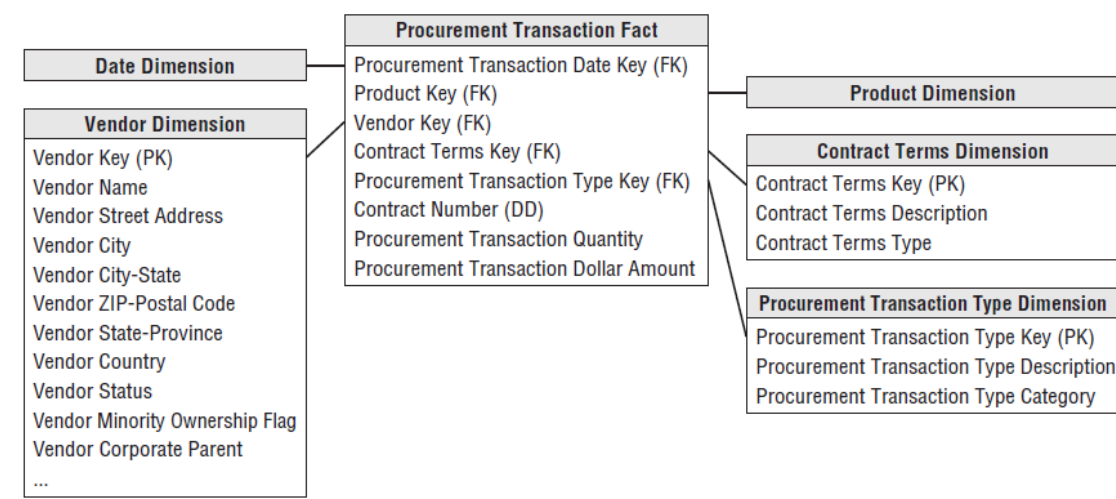
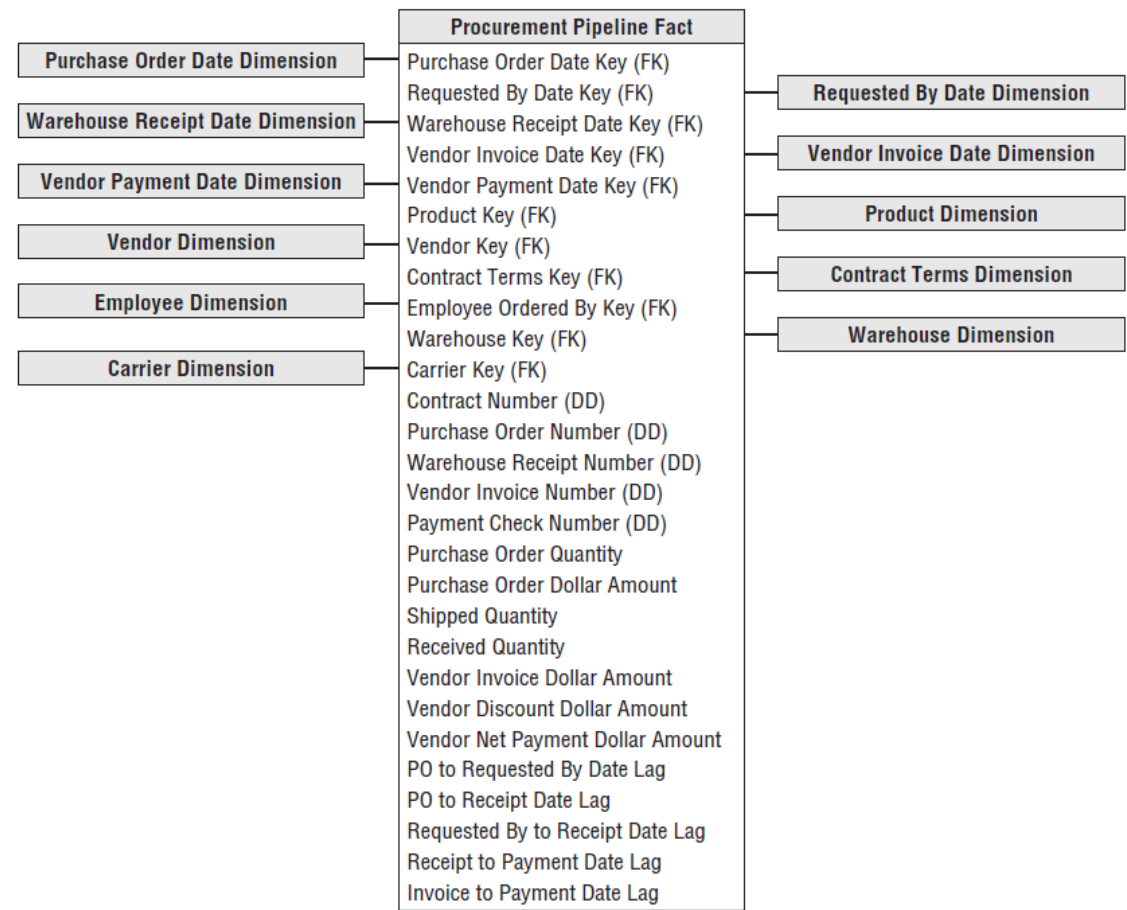
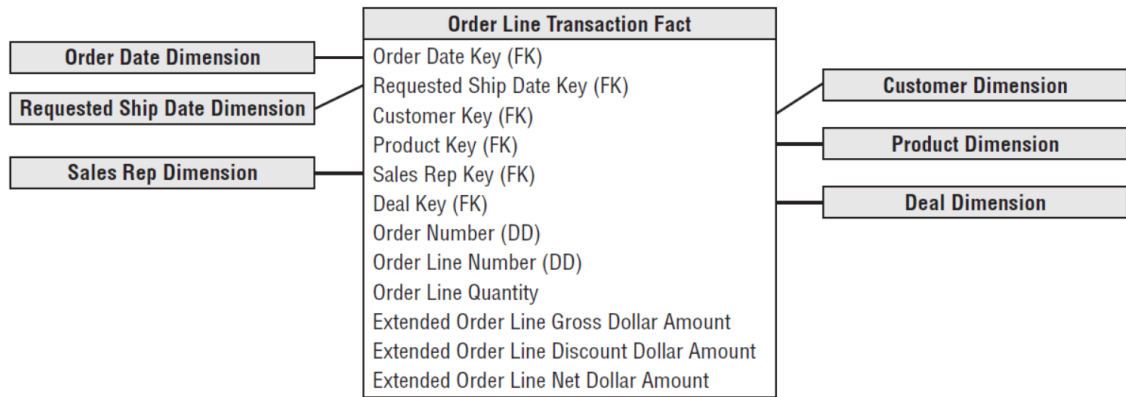




Schéma 12



## Schéma 13



## Schéma 14

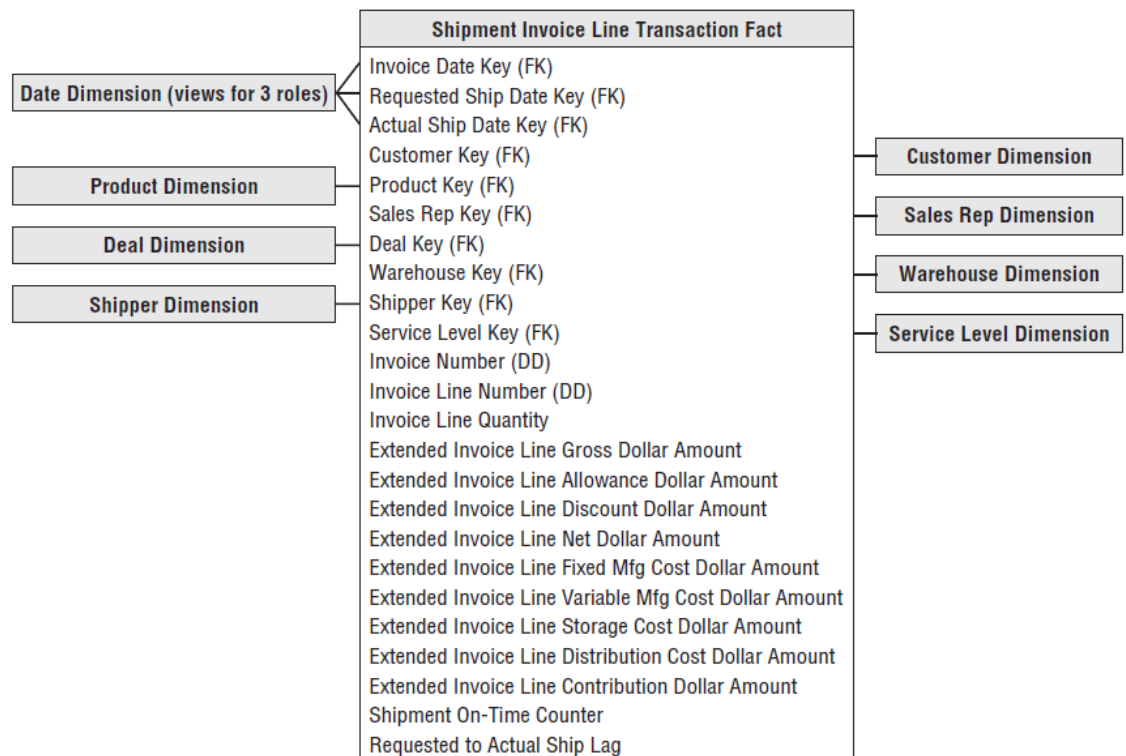


Schéma 15

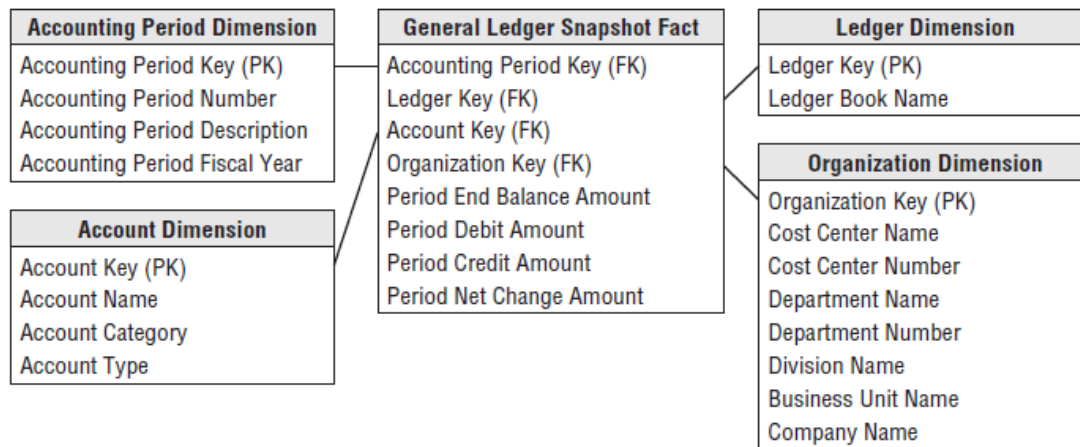


Schéma 16

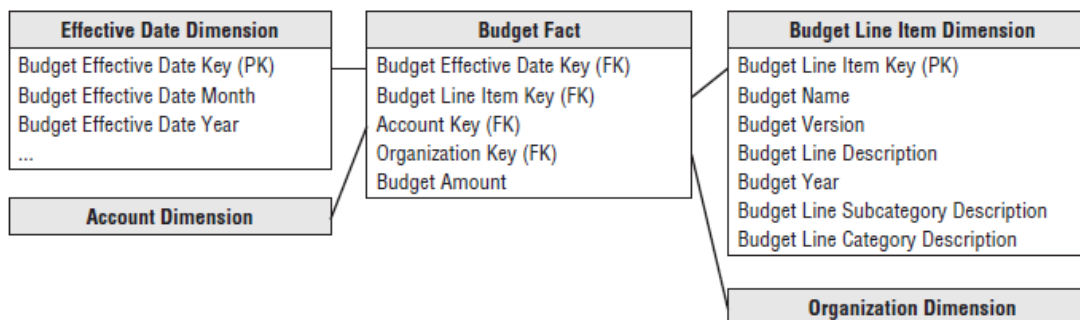


Schéma 17

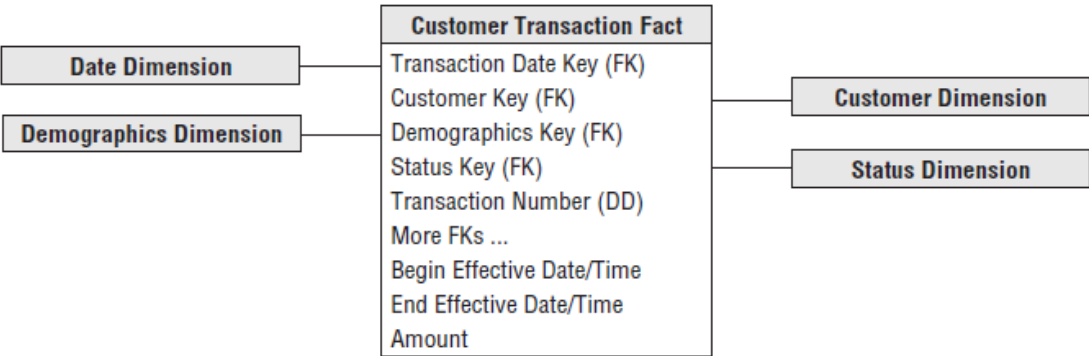
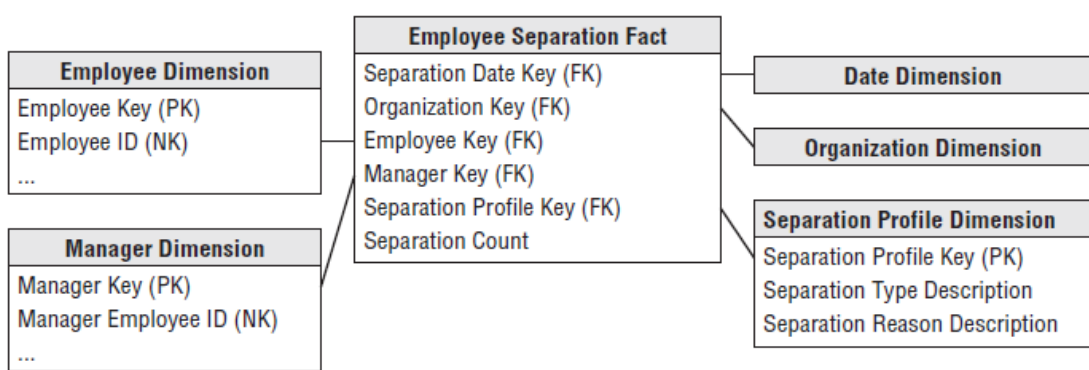


Schéma 18



## Schéma 19

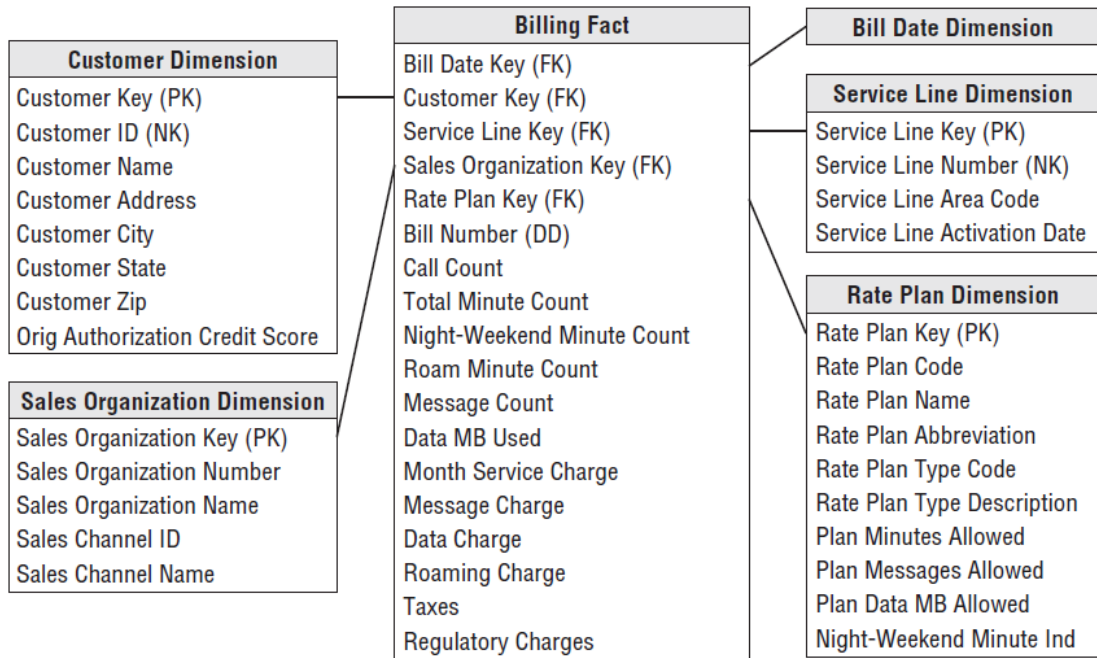


Schéma 20

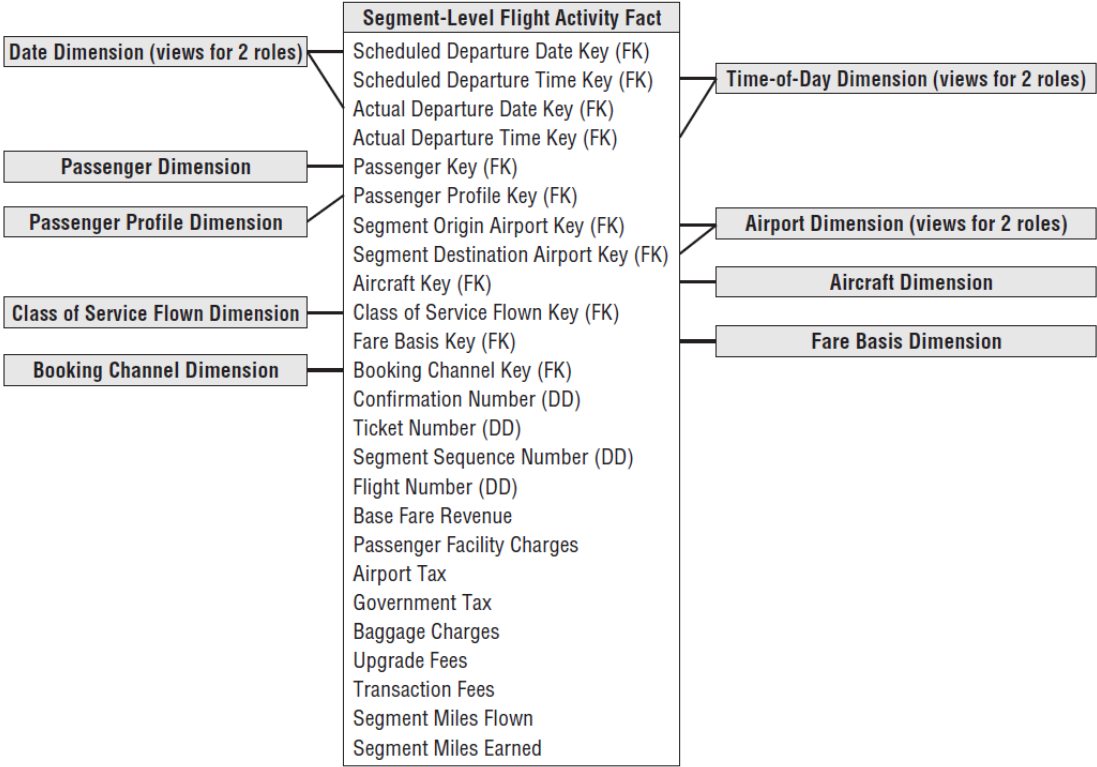


Schéma 21

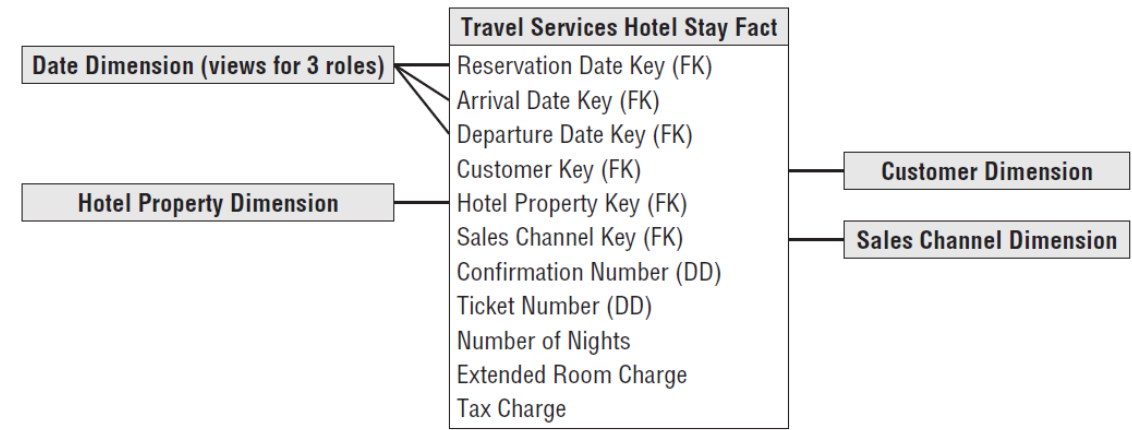
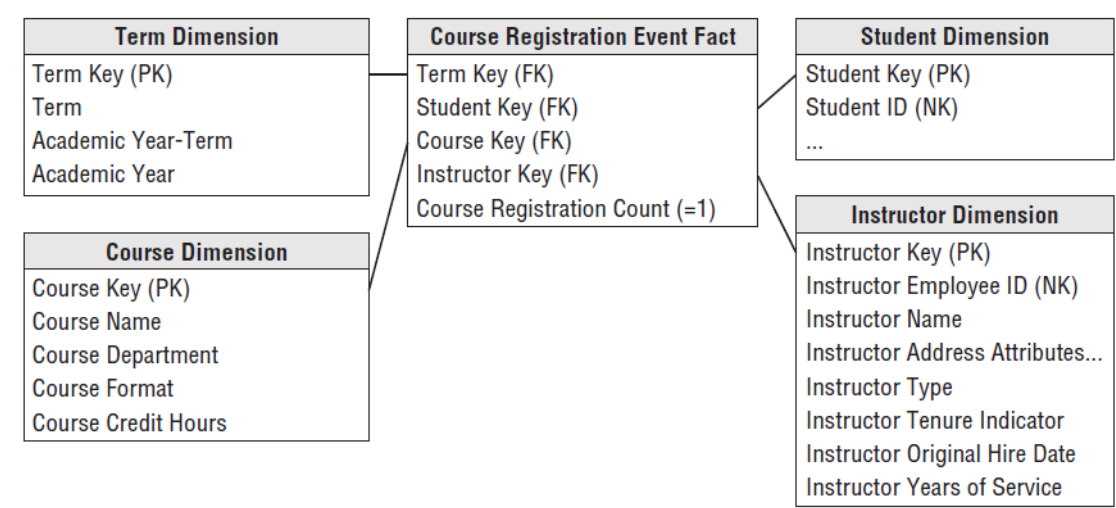
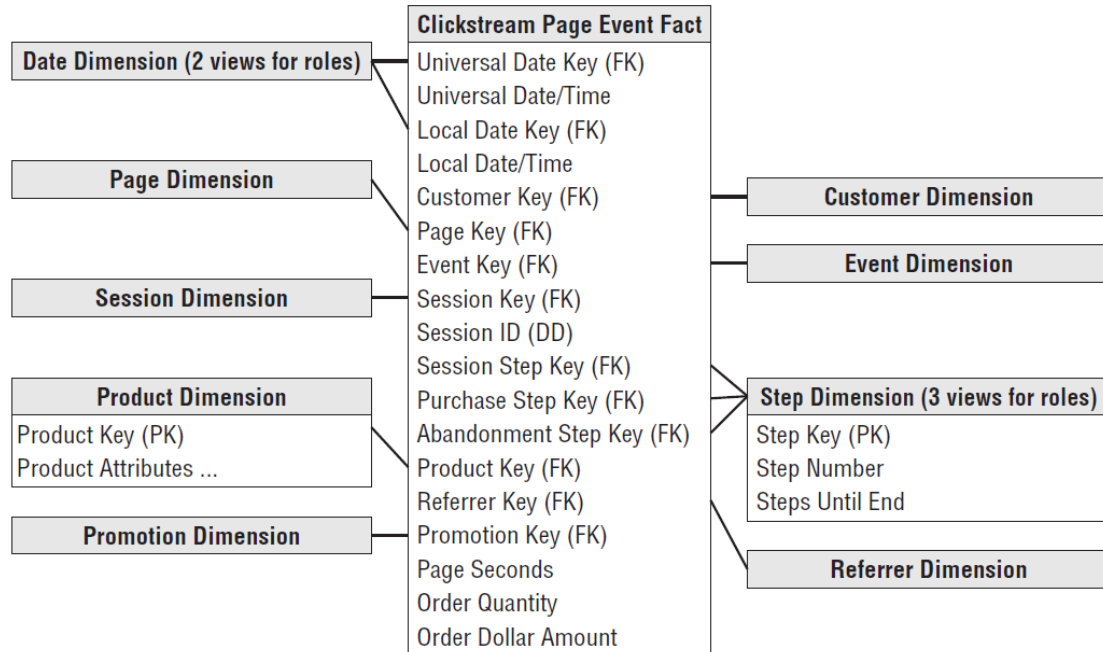


Schéma 22



## Schéma 23



## Schéma 24

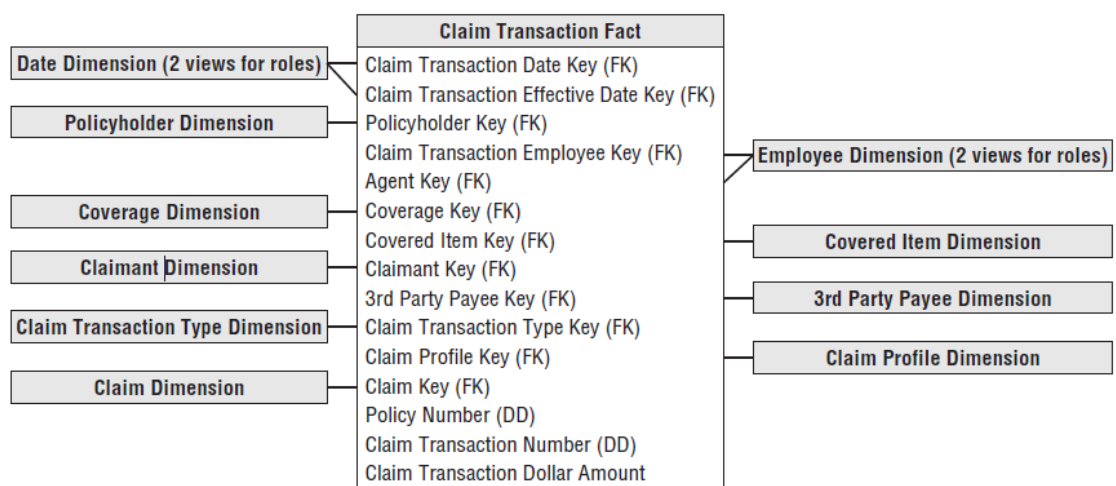




Schéma 25

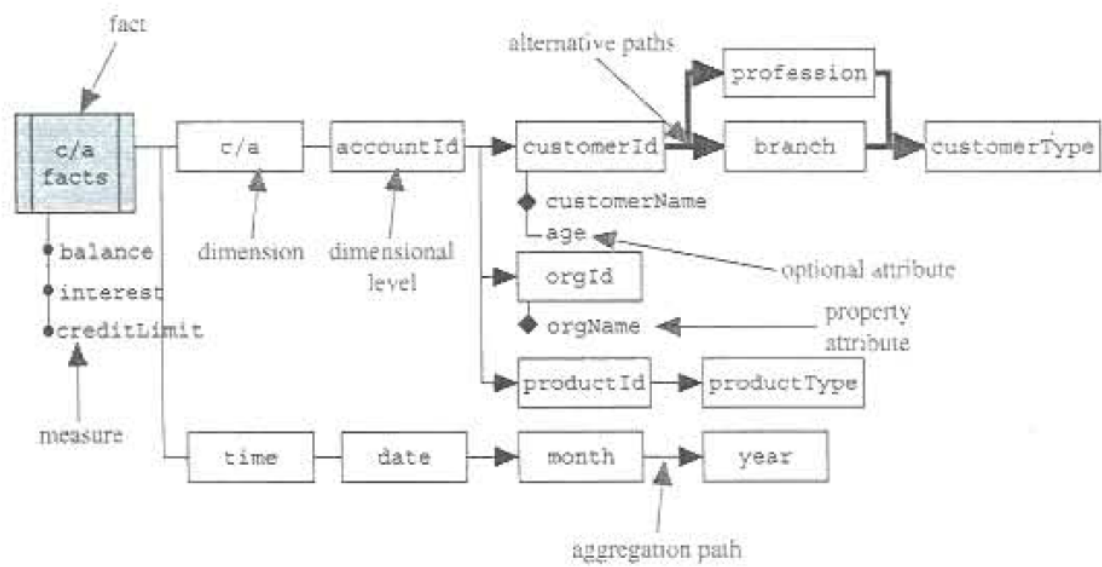


Schéma 26

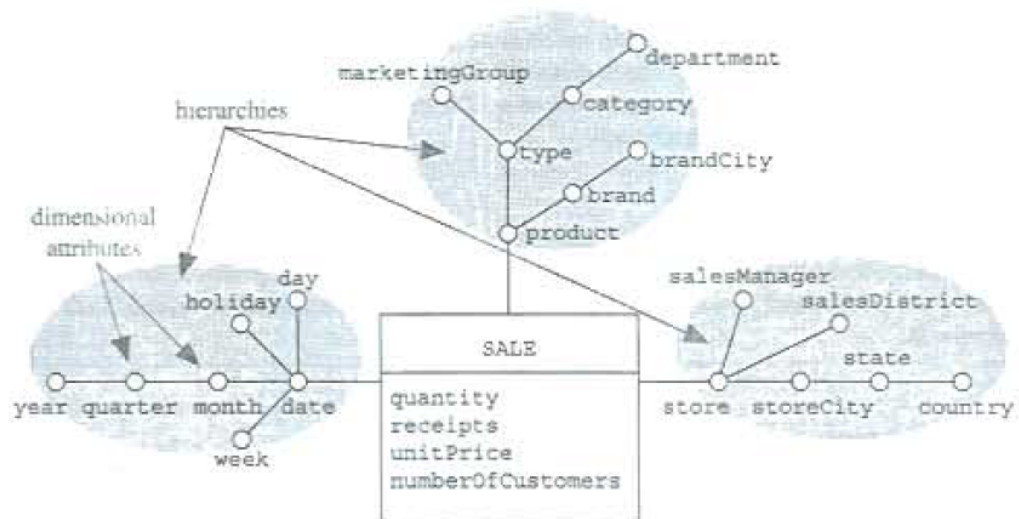


Schéma 27

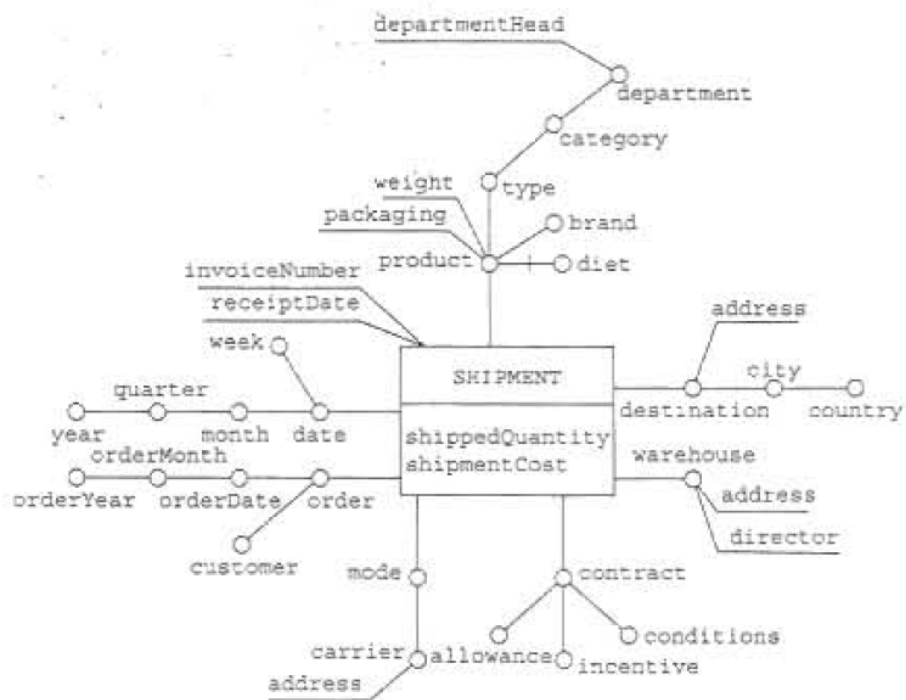


Schéma 28

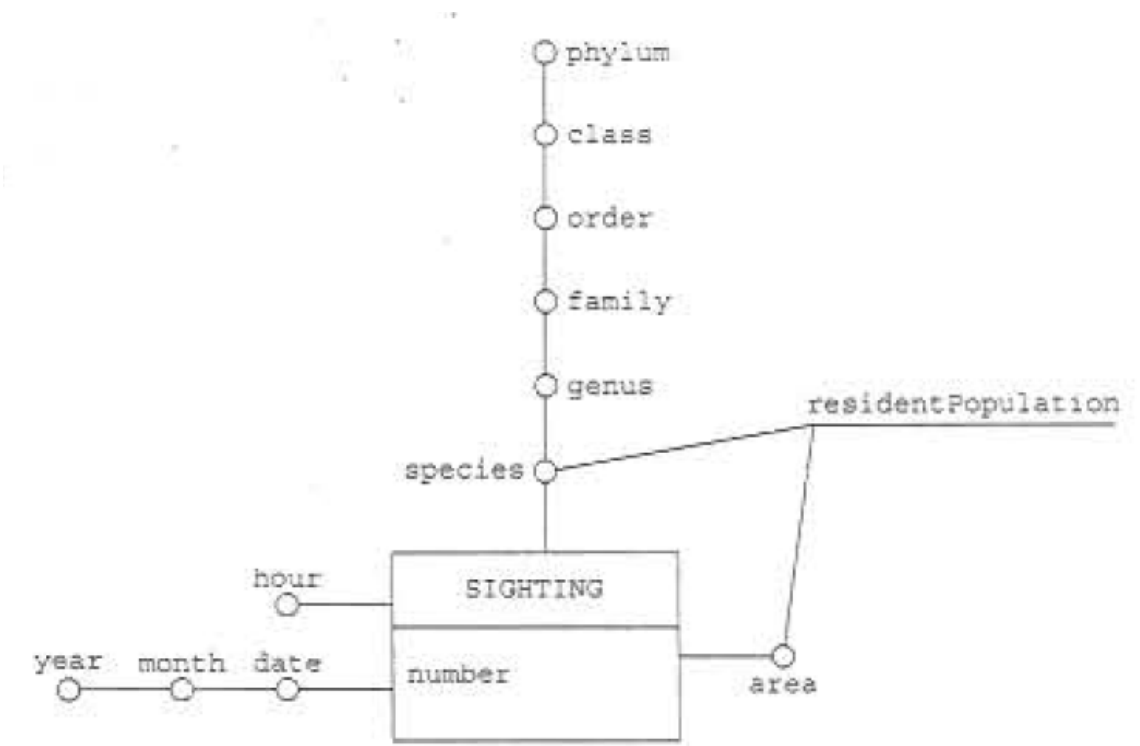


Schéma 29

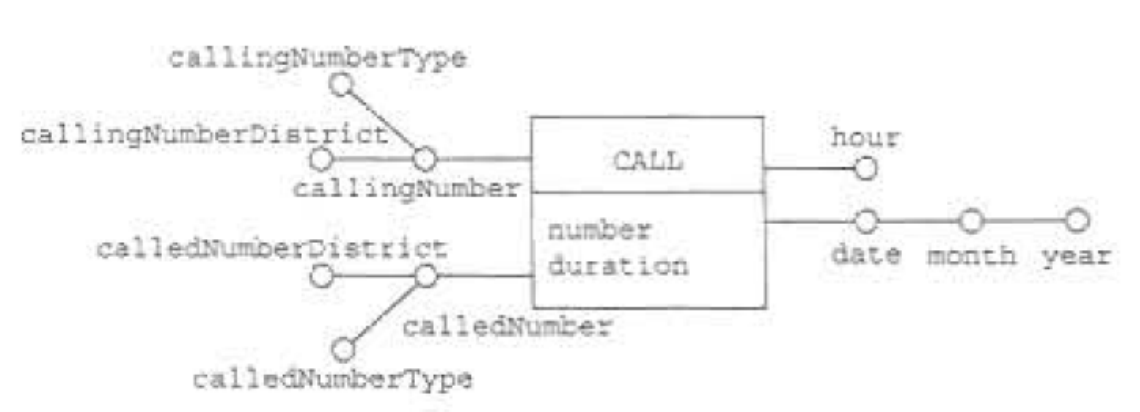


Schéma 30

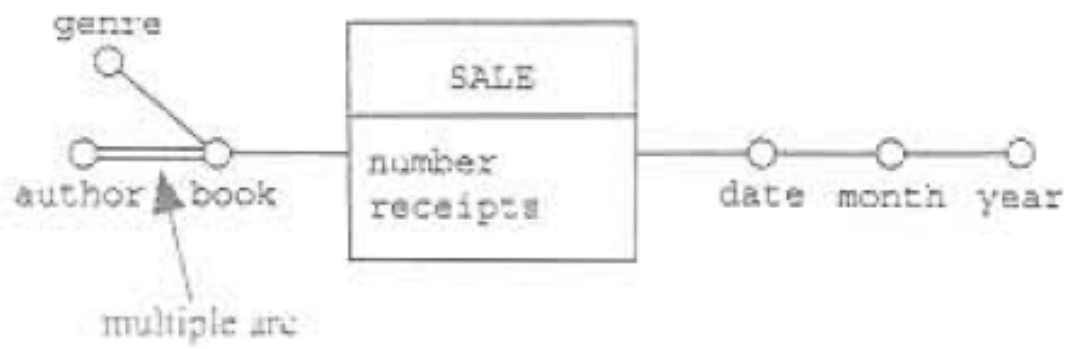


Schéma 31

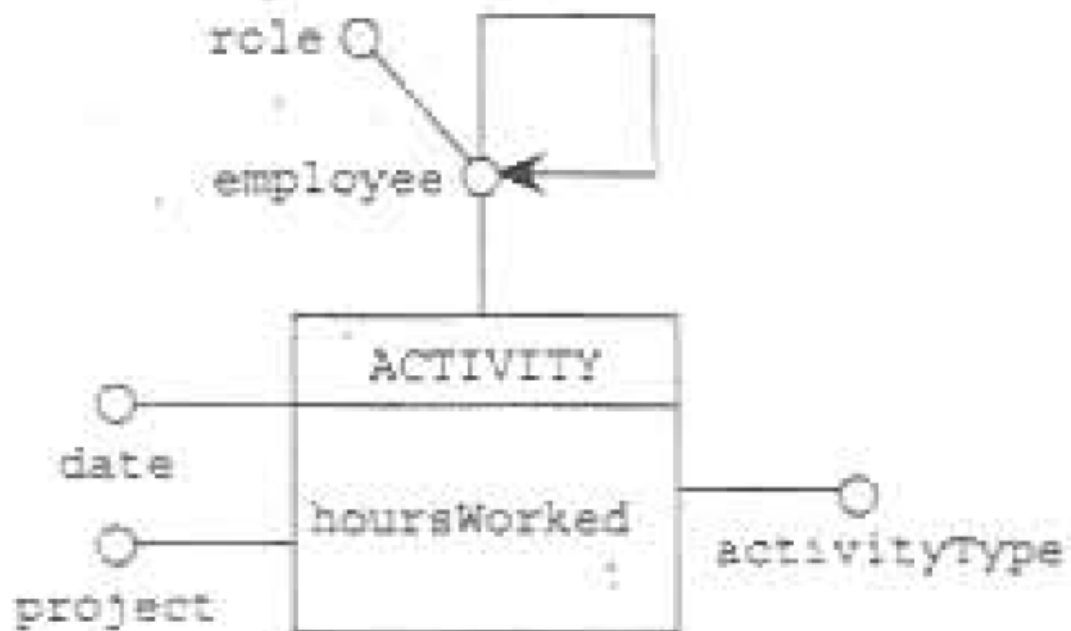


Schéma 32

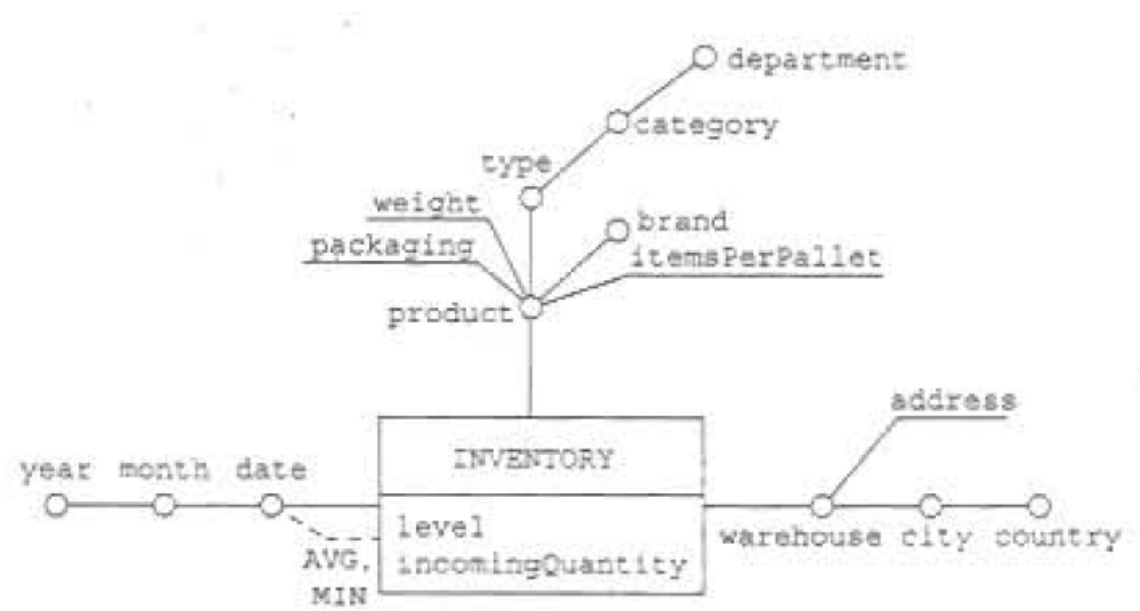


Schéma 33

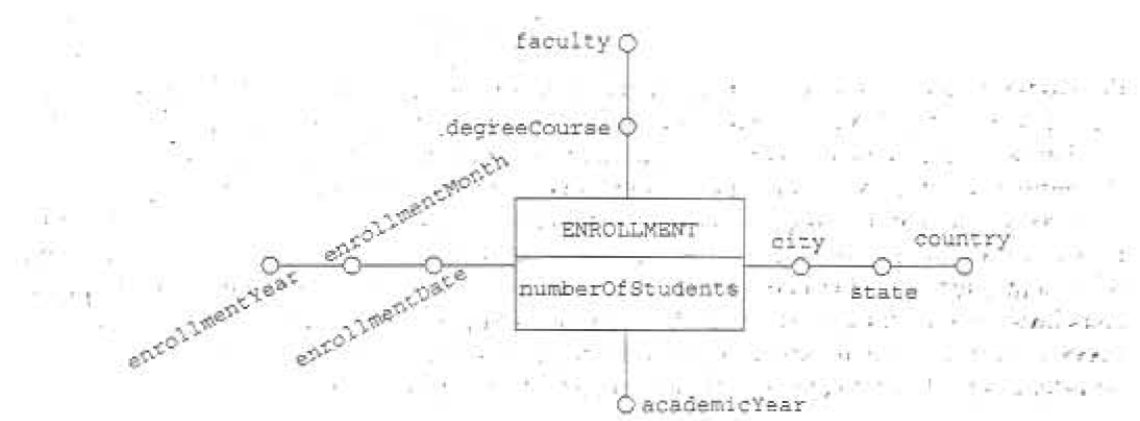


Schéma 34

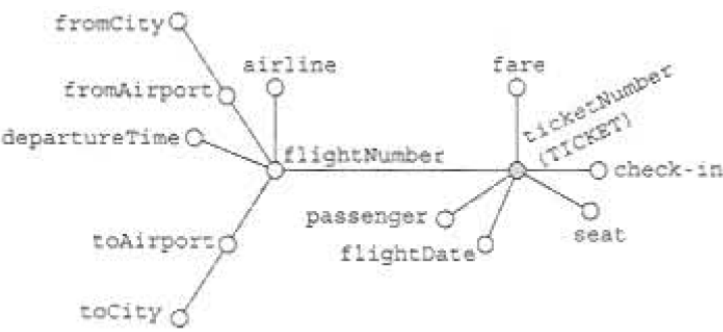


Schéma 35

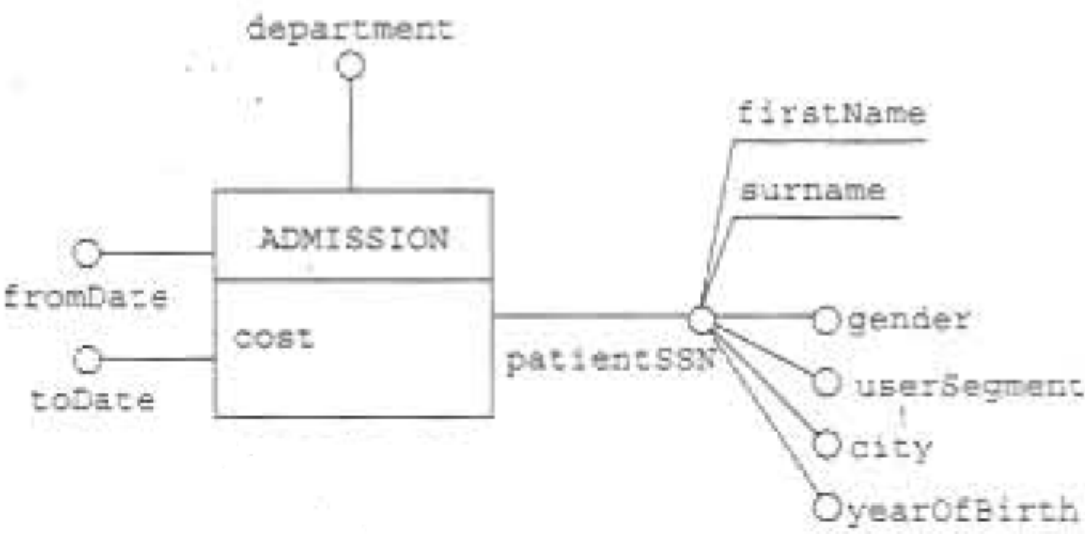
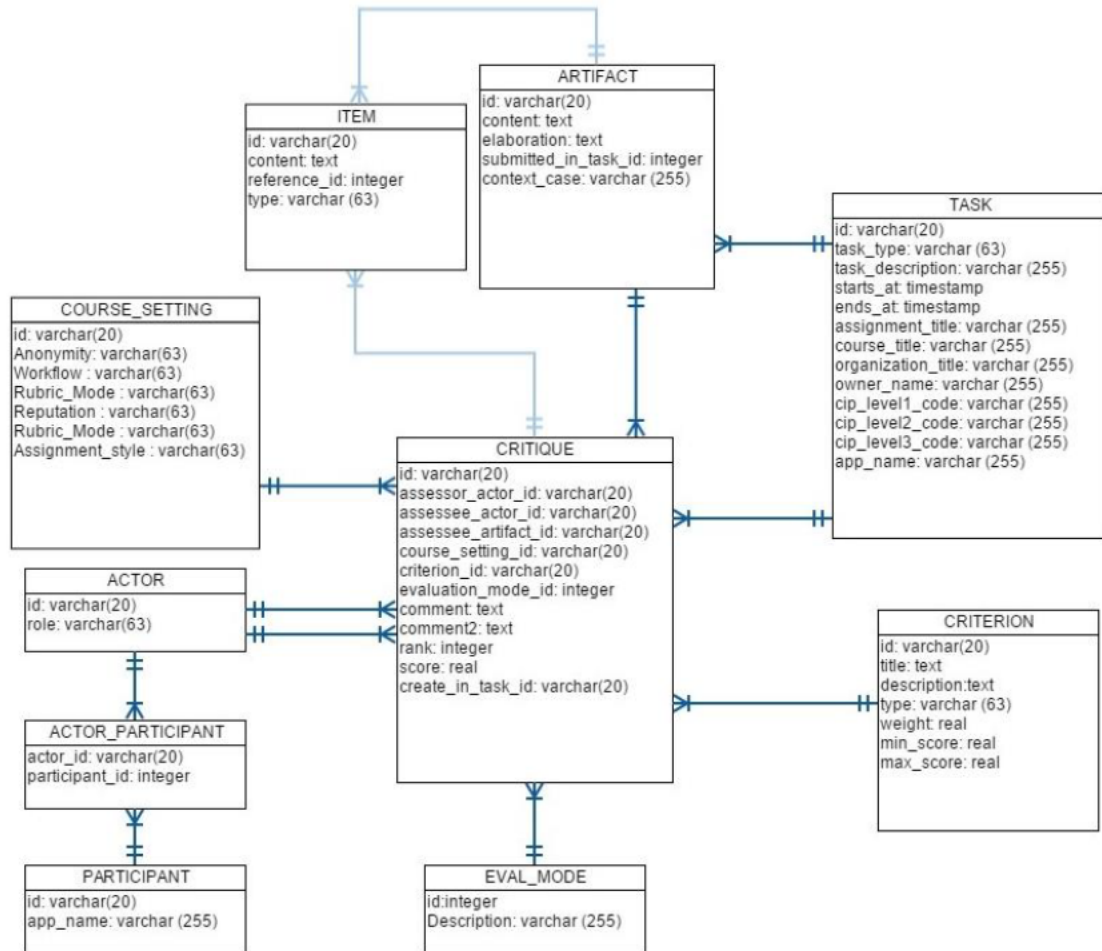


Schéma 36



## Schéma 37

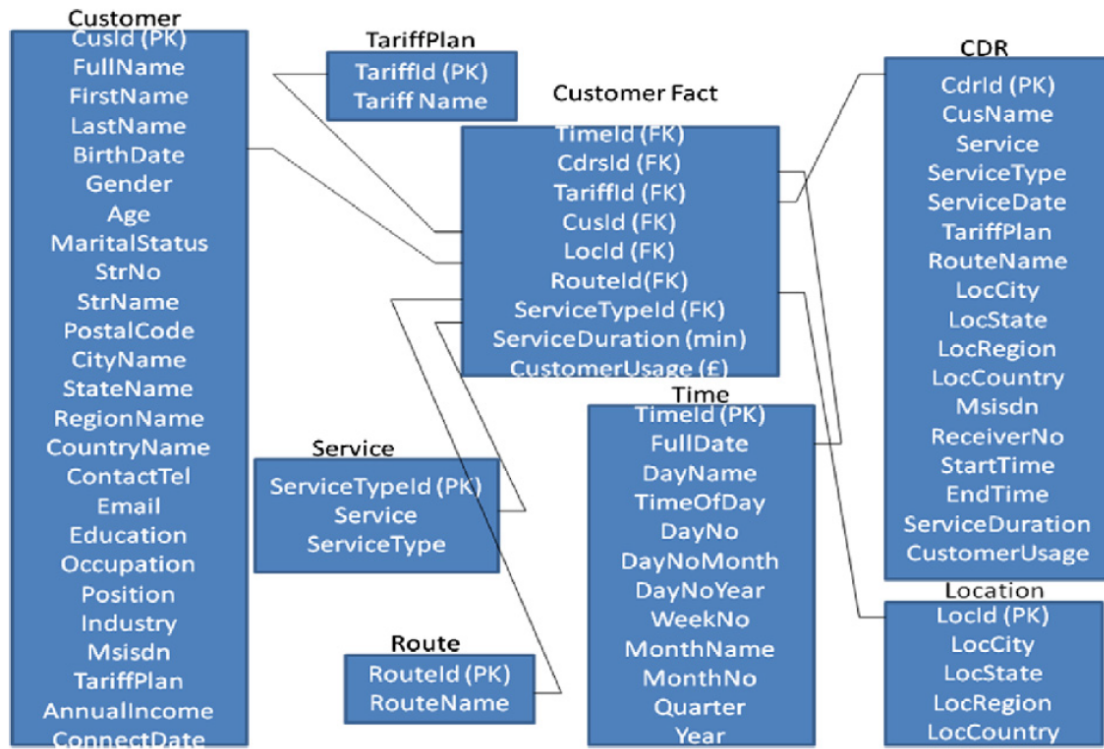
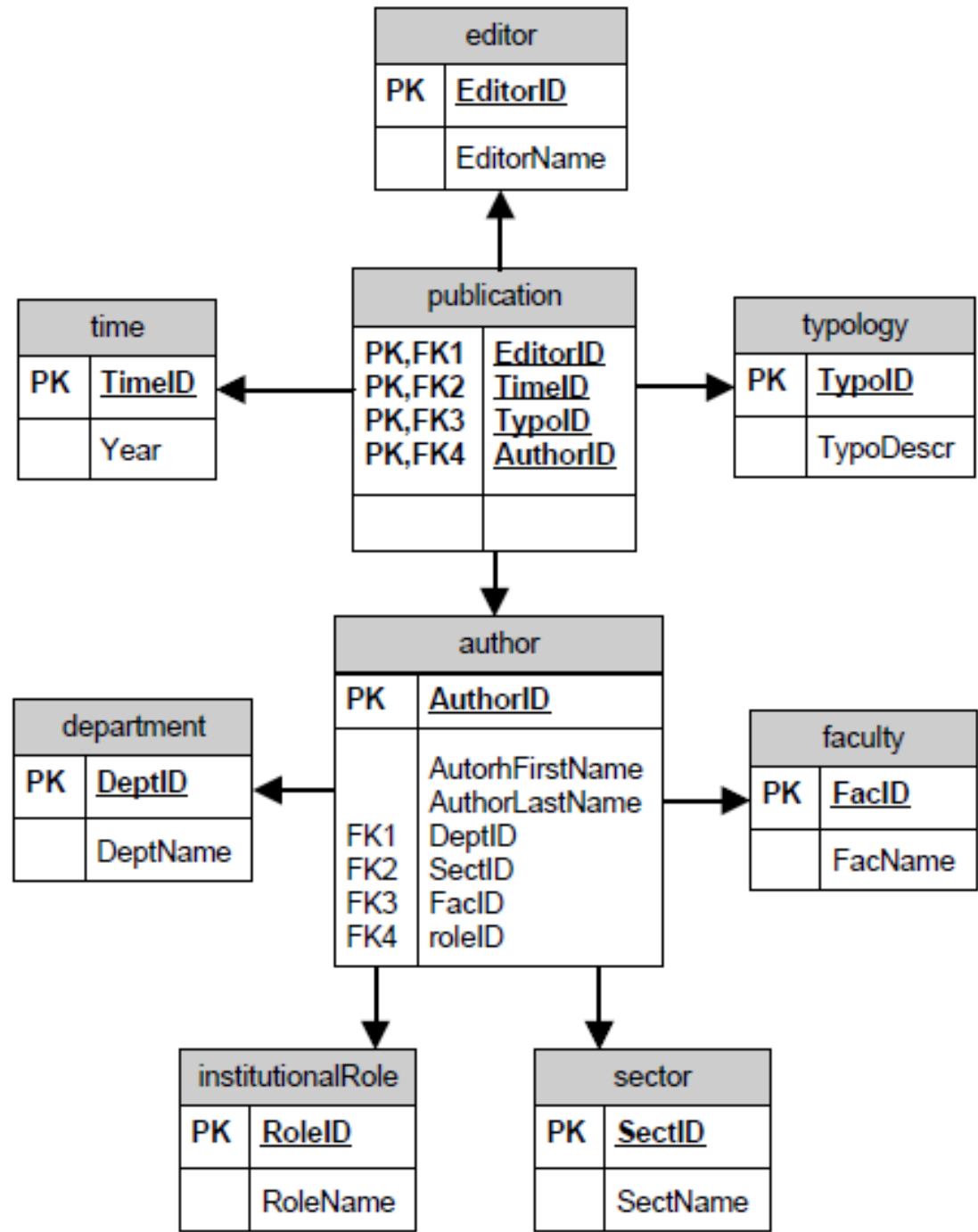




Schéma 38



## Schéma 39

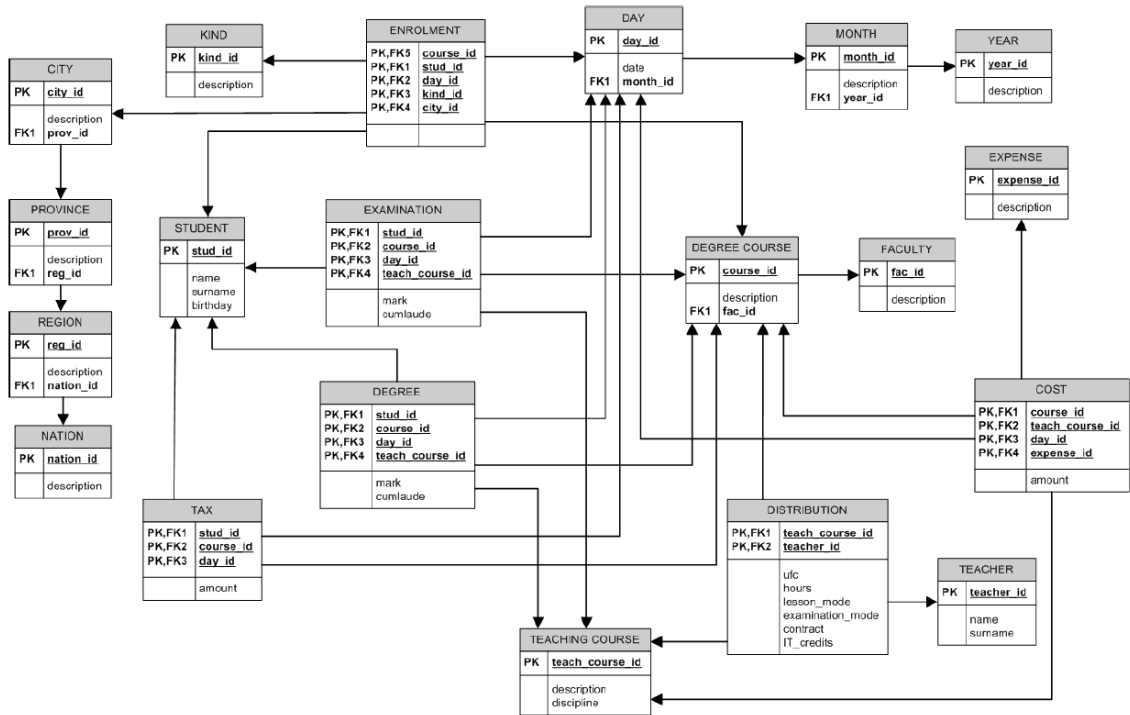
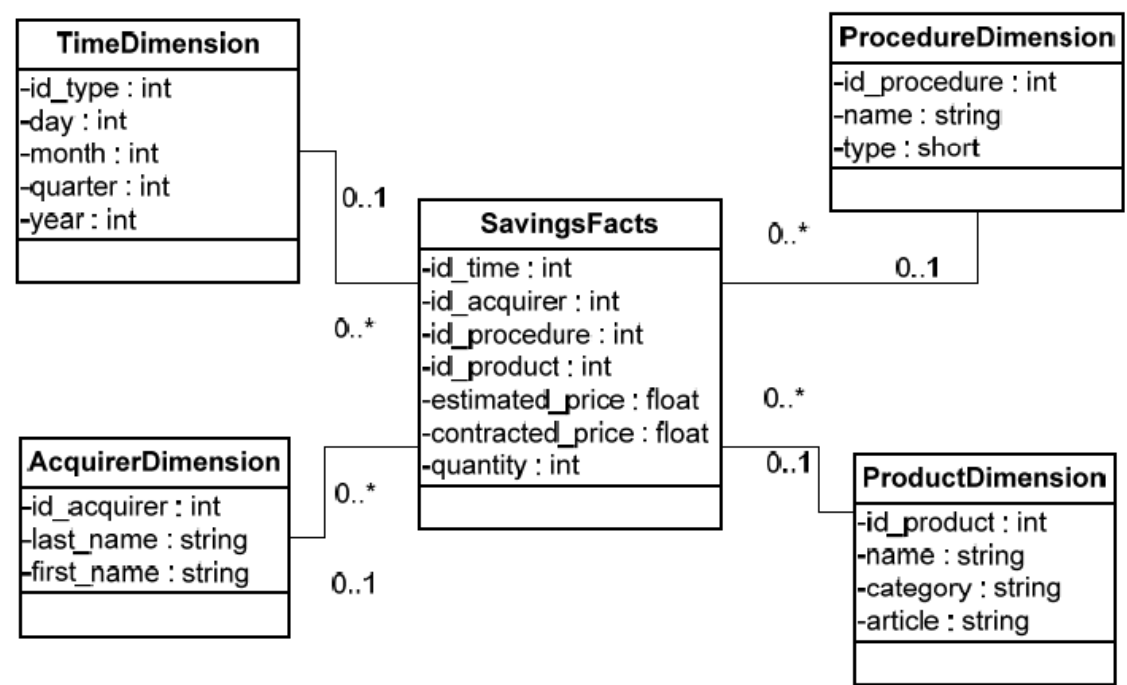


Schéma 40



## Schéma 41

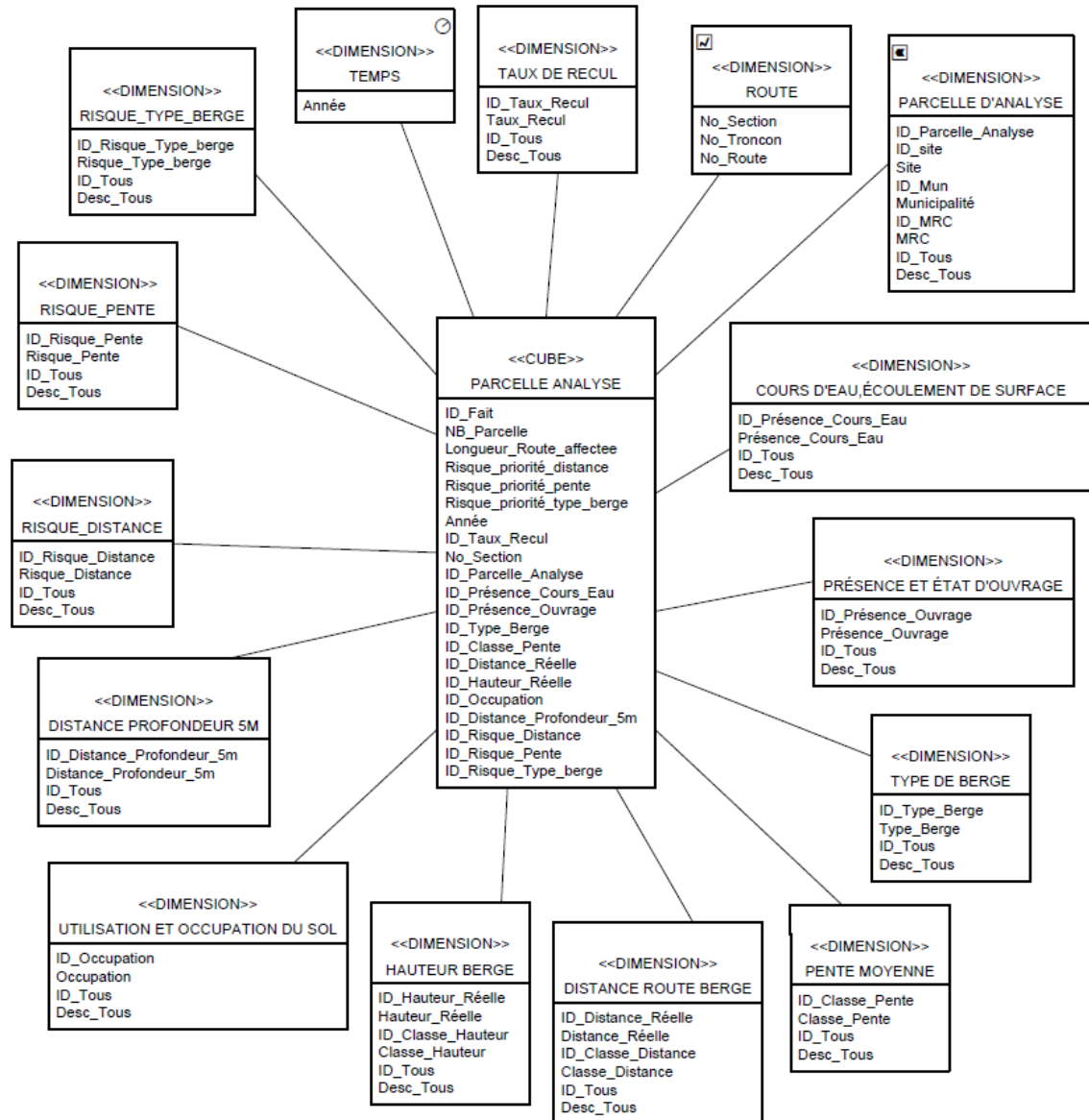
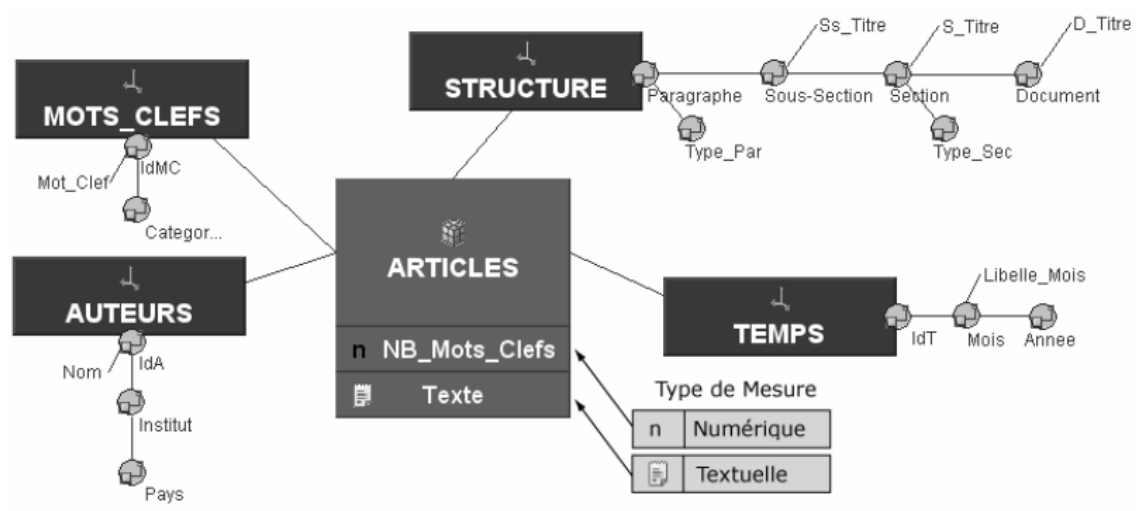


Schéma 42



## Schéma 43

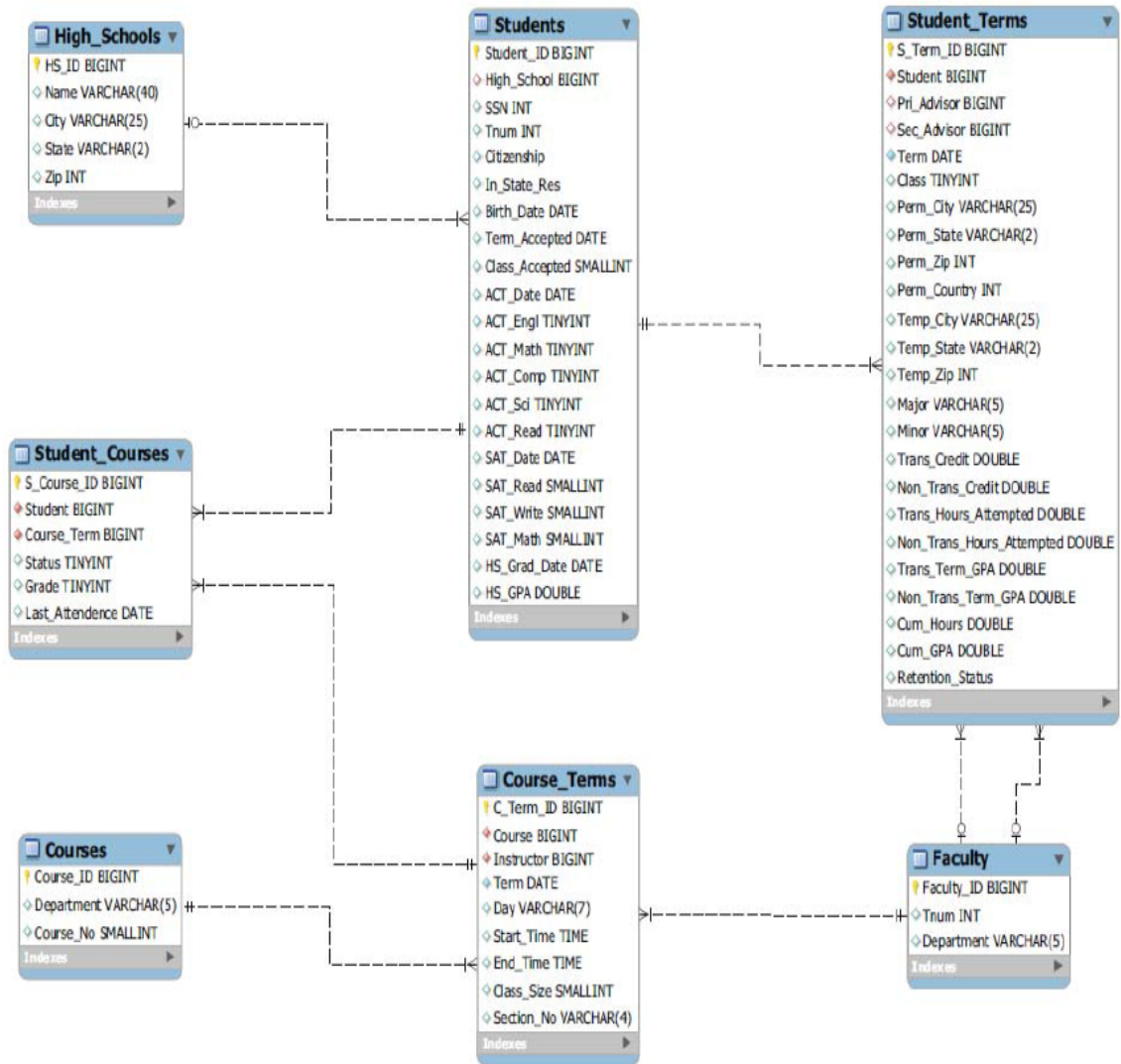


Schéma 44

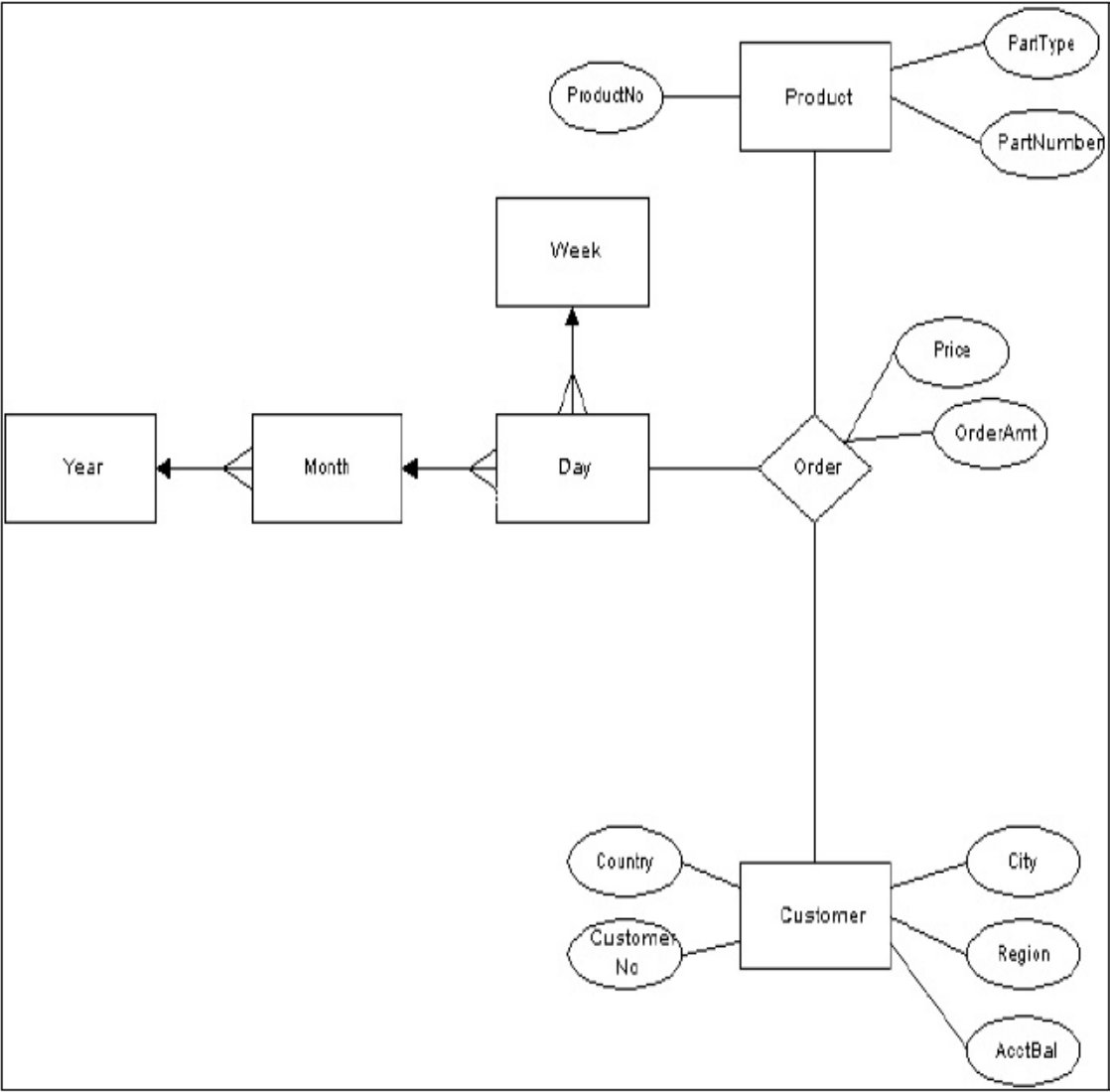


Schéma 45

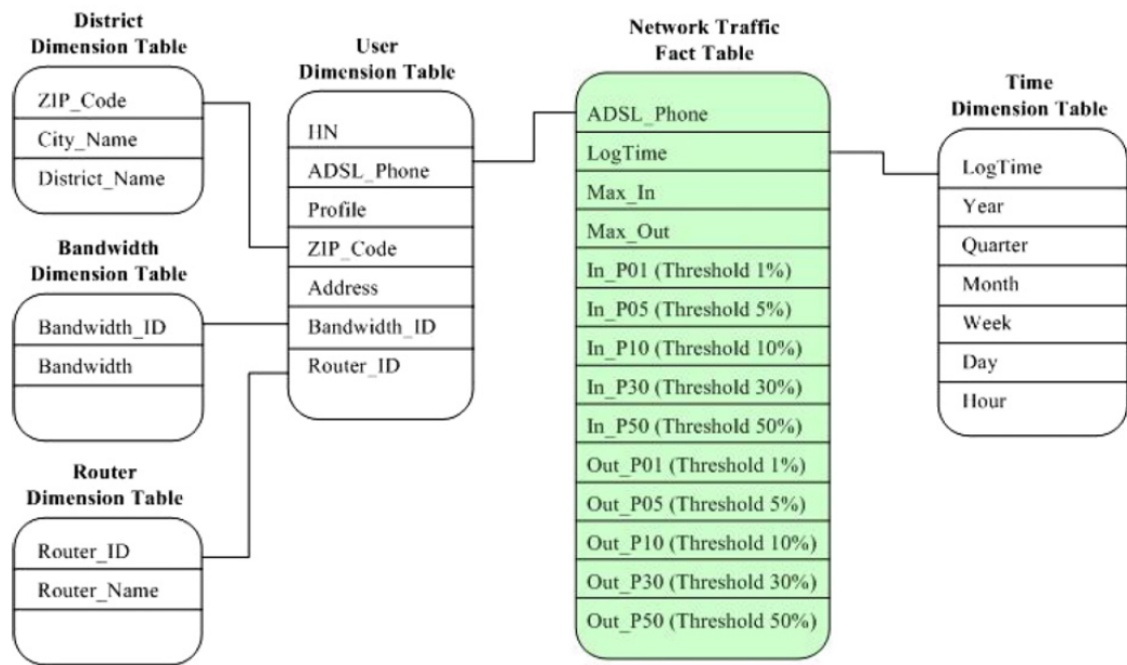
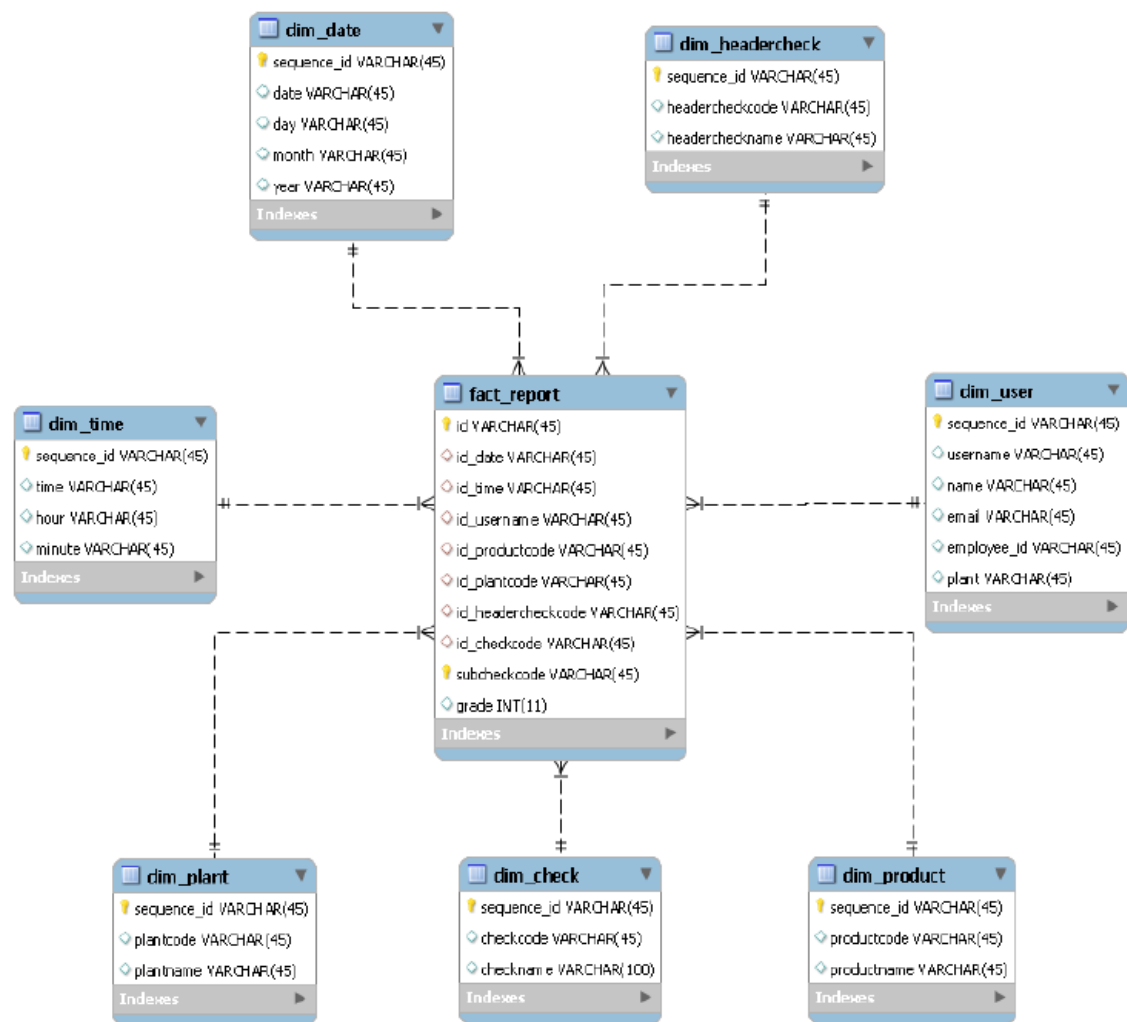




Schéma 46



## Schéma 47

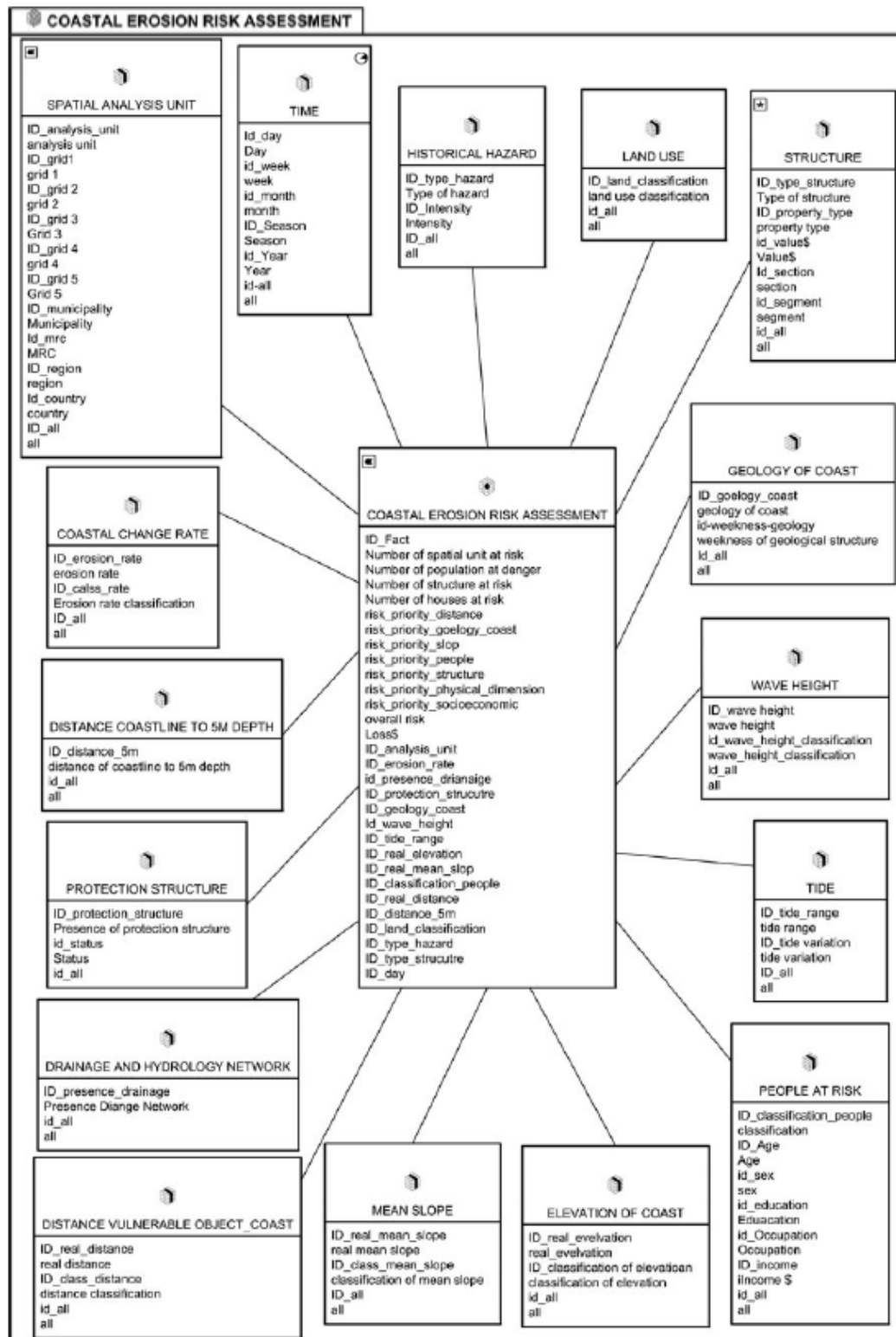


Schéma 48

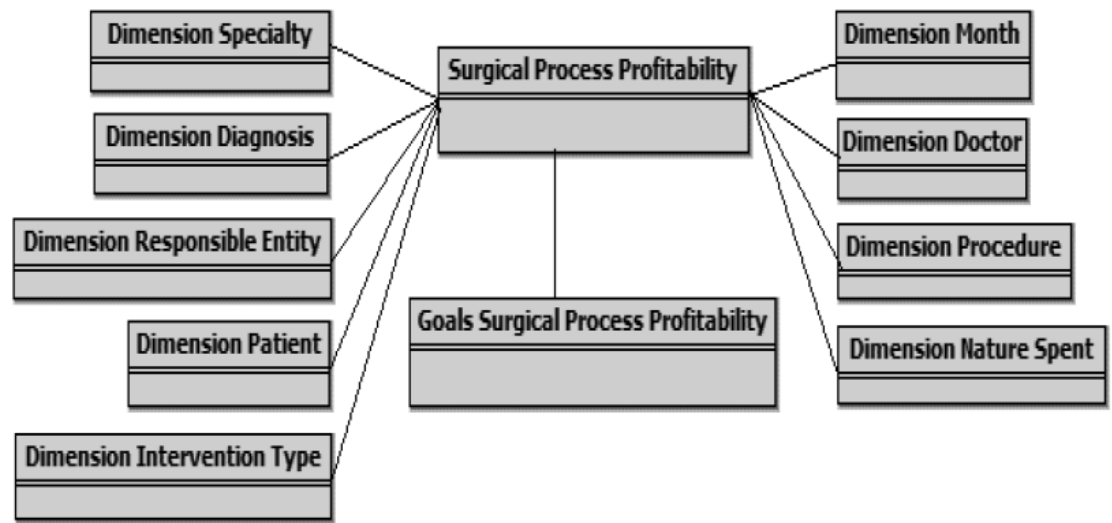


Schéma 49

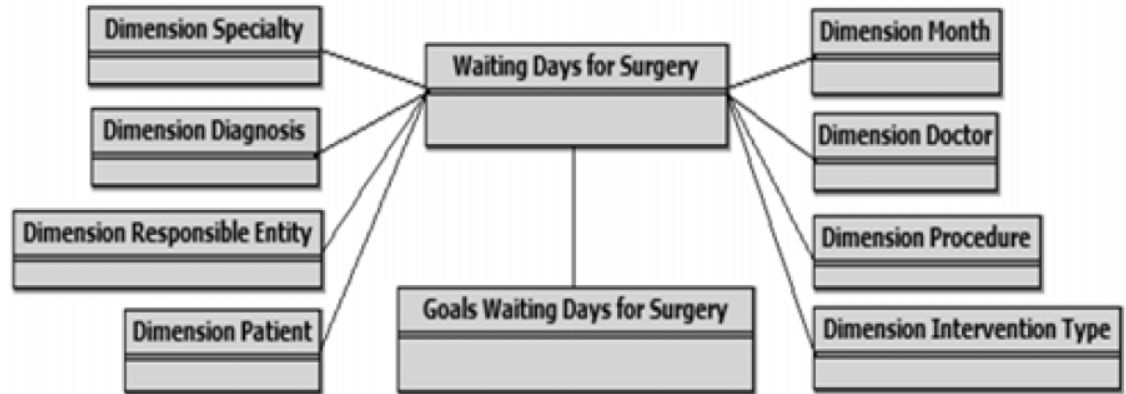


Schéma 50

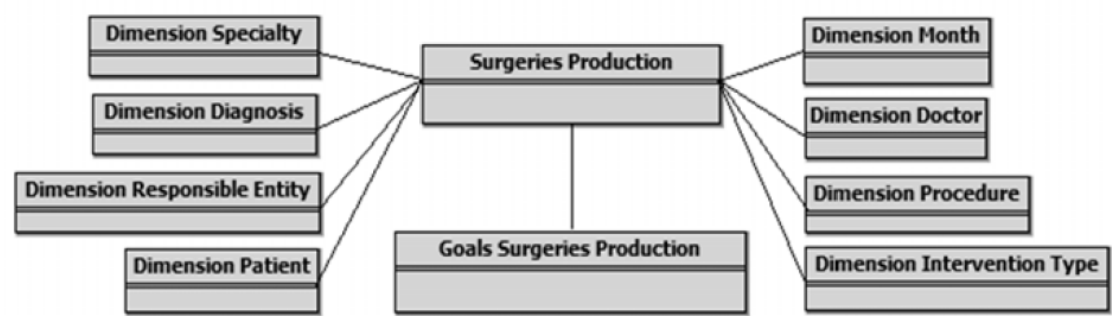
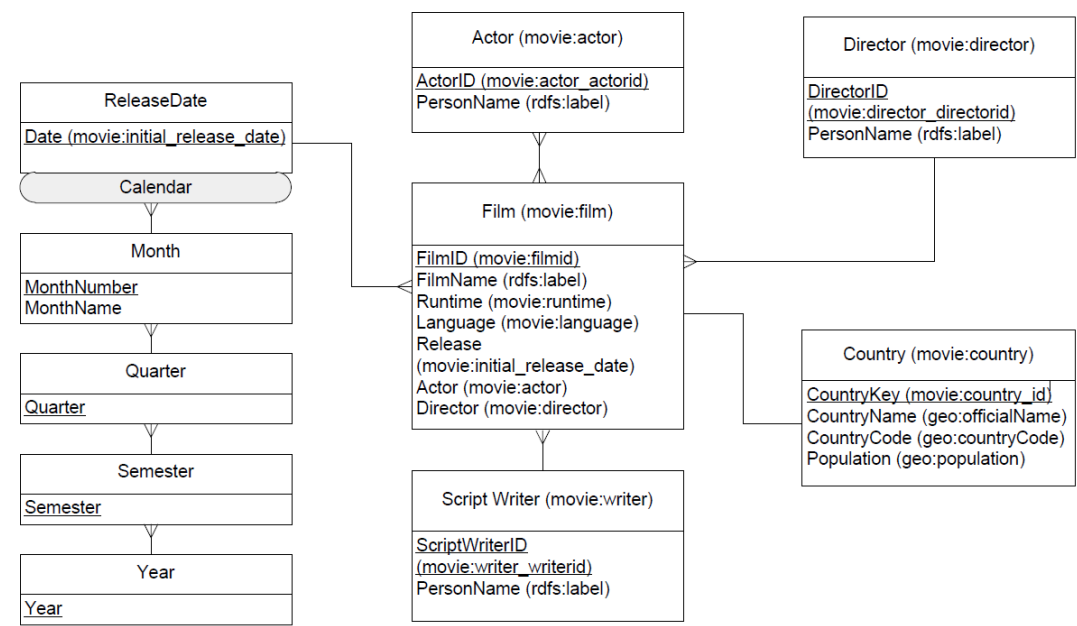


Schéma 51



## Schéma 52

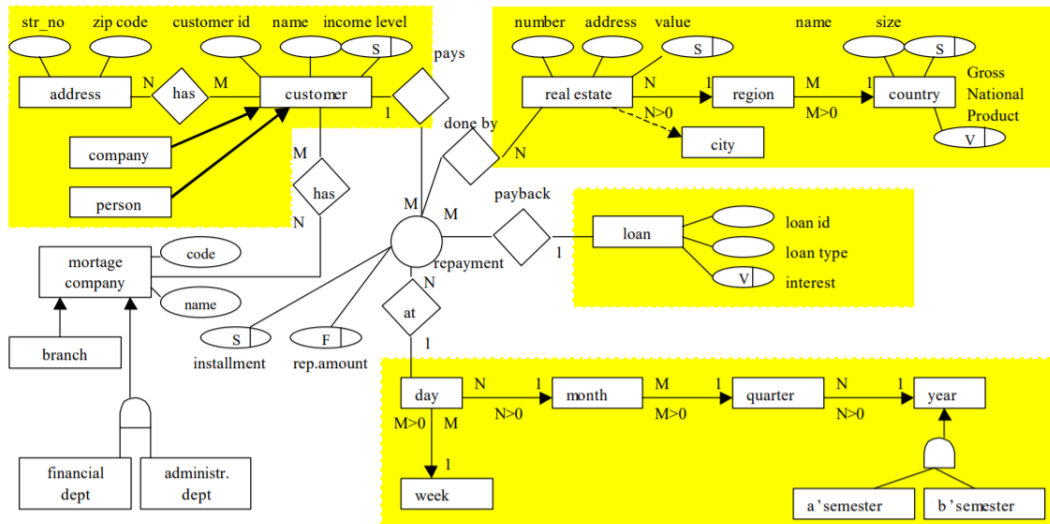


Schéma 53

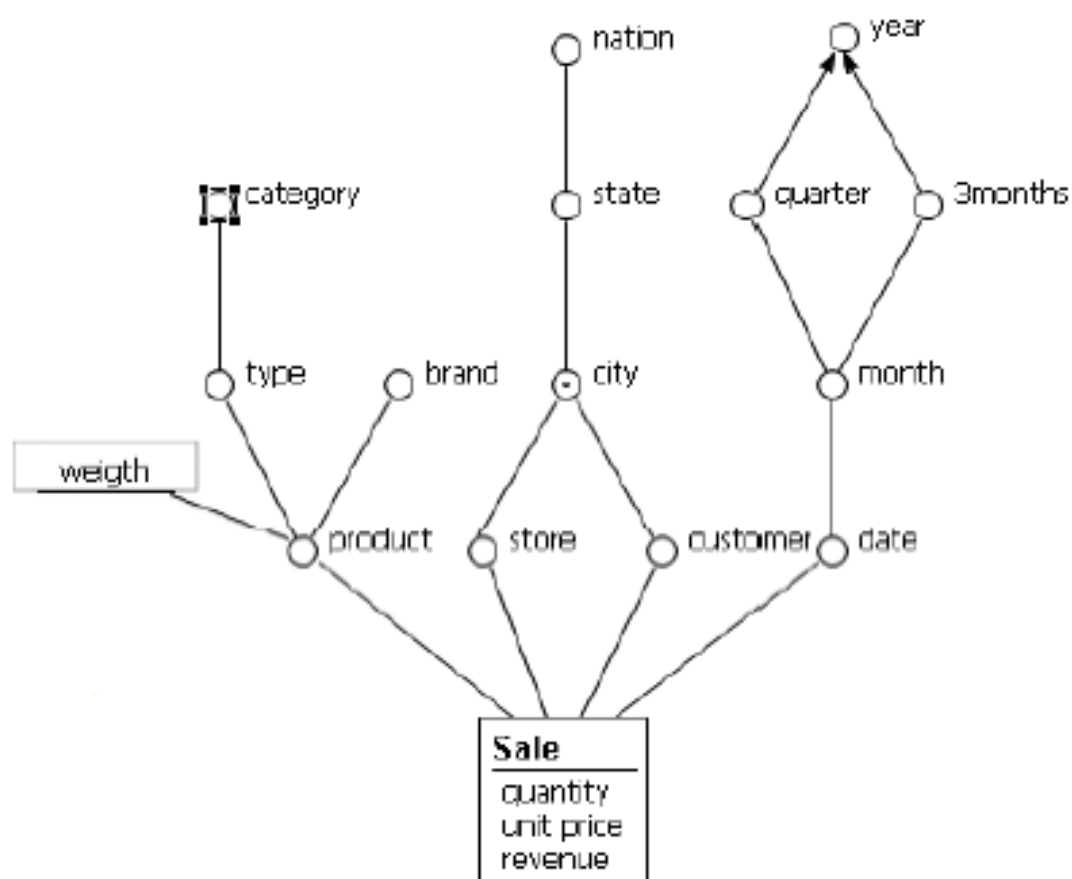


Schéma 54

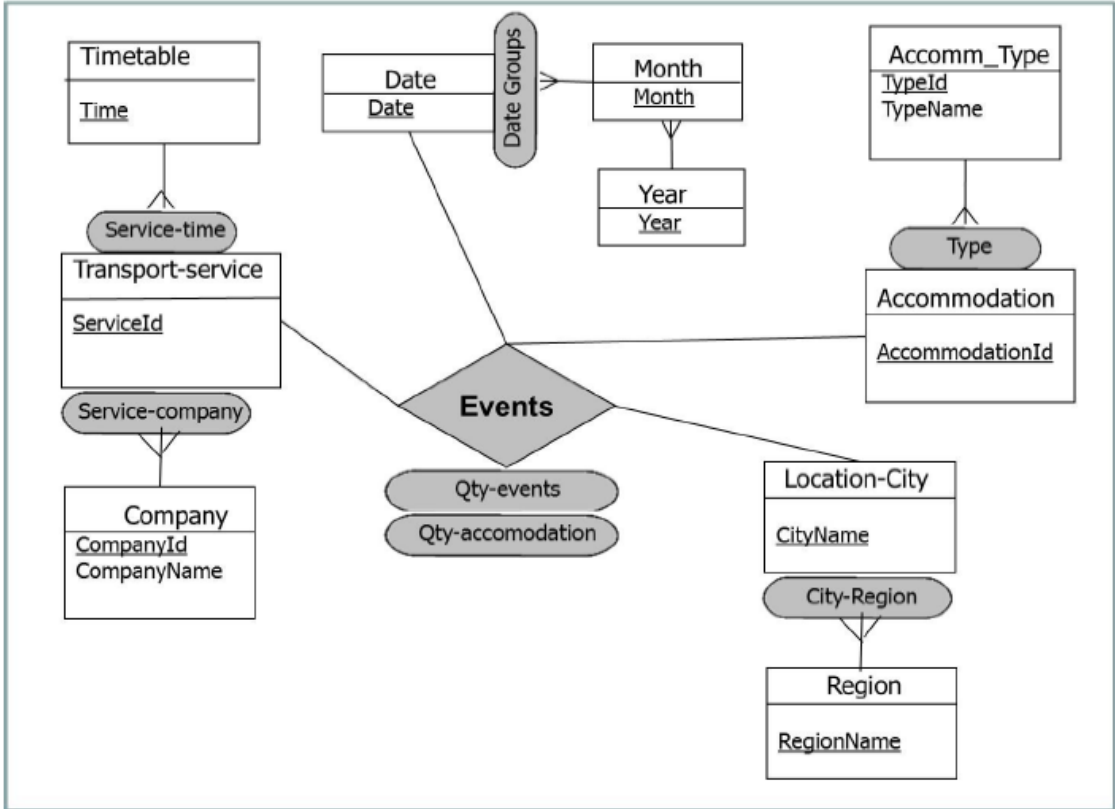


Schéma 55

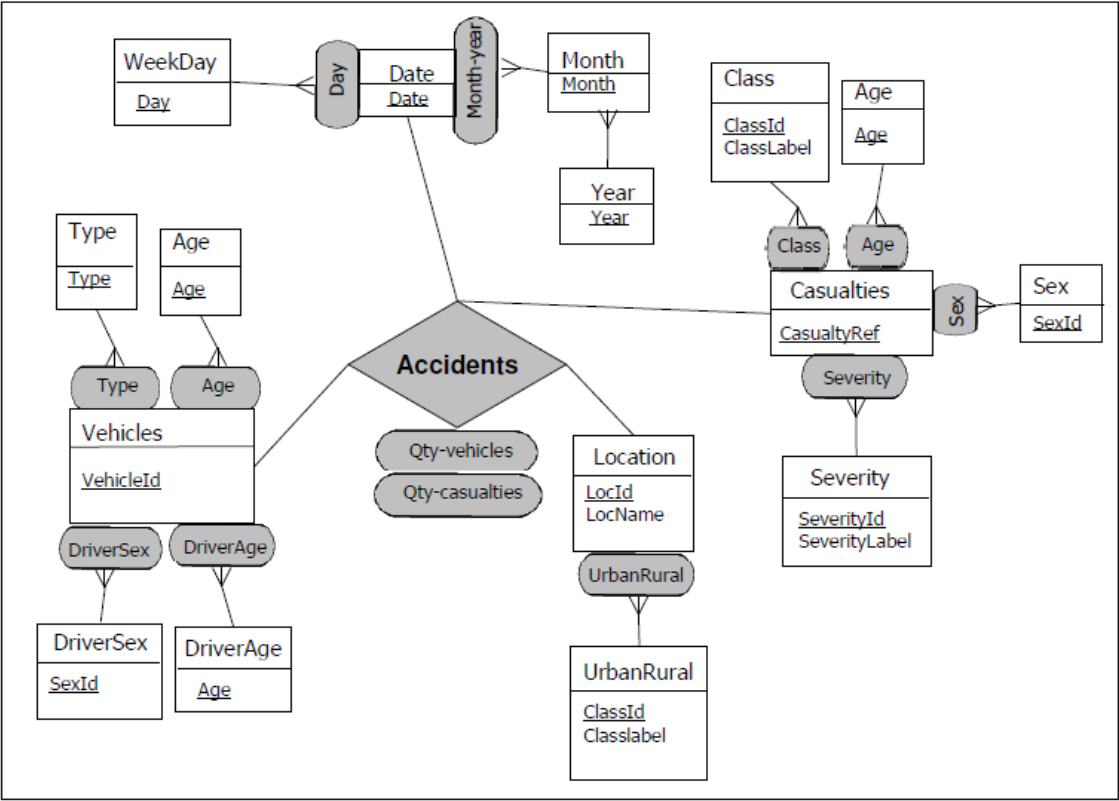




Schéma 56

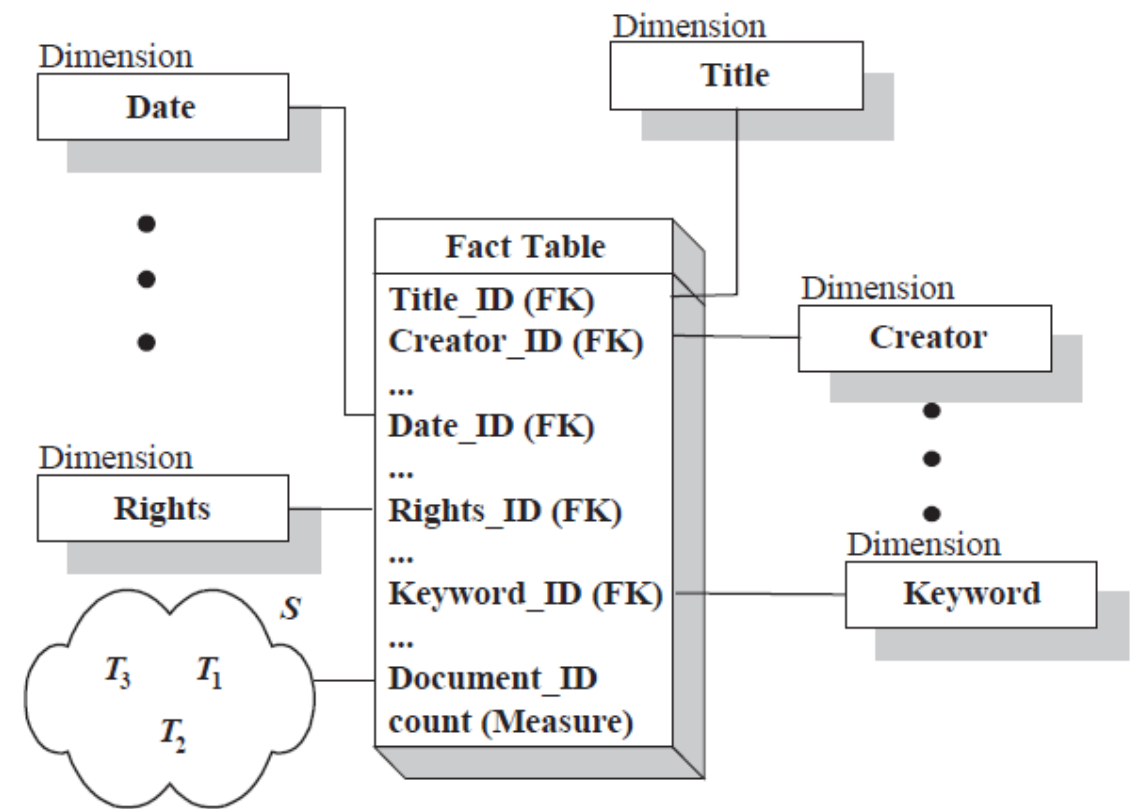


Schéma 57

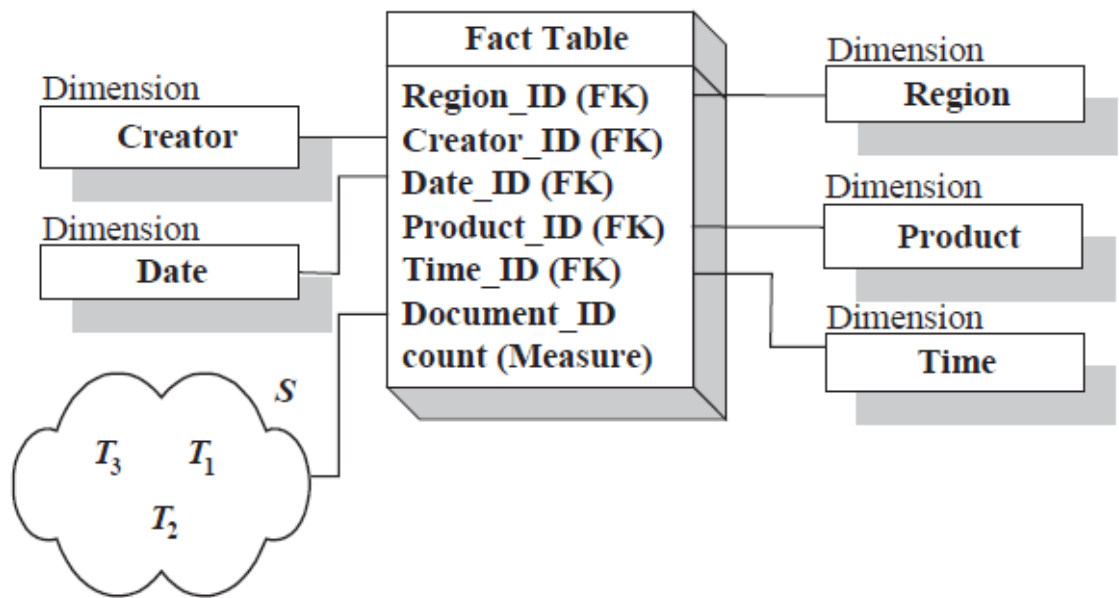


Schéma 58

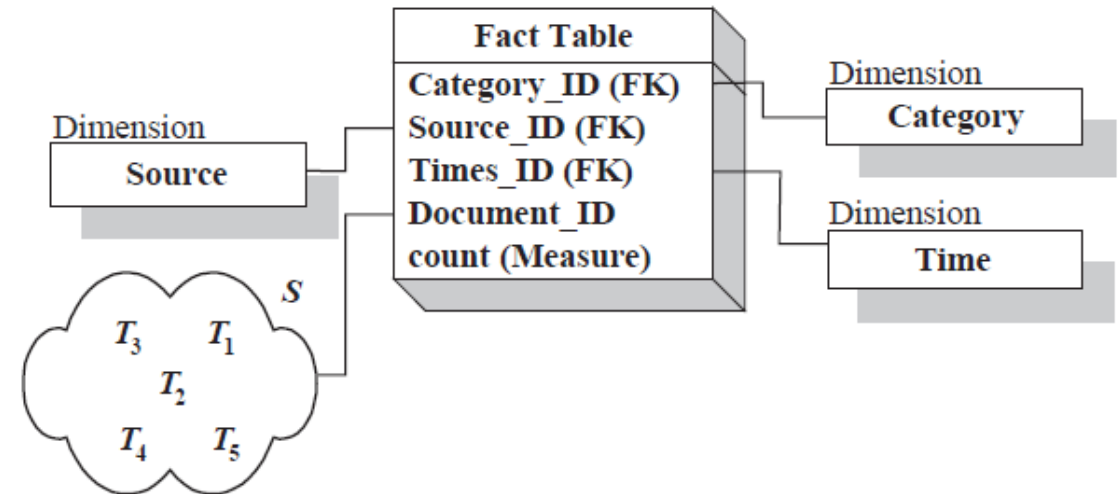


Schéma 59

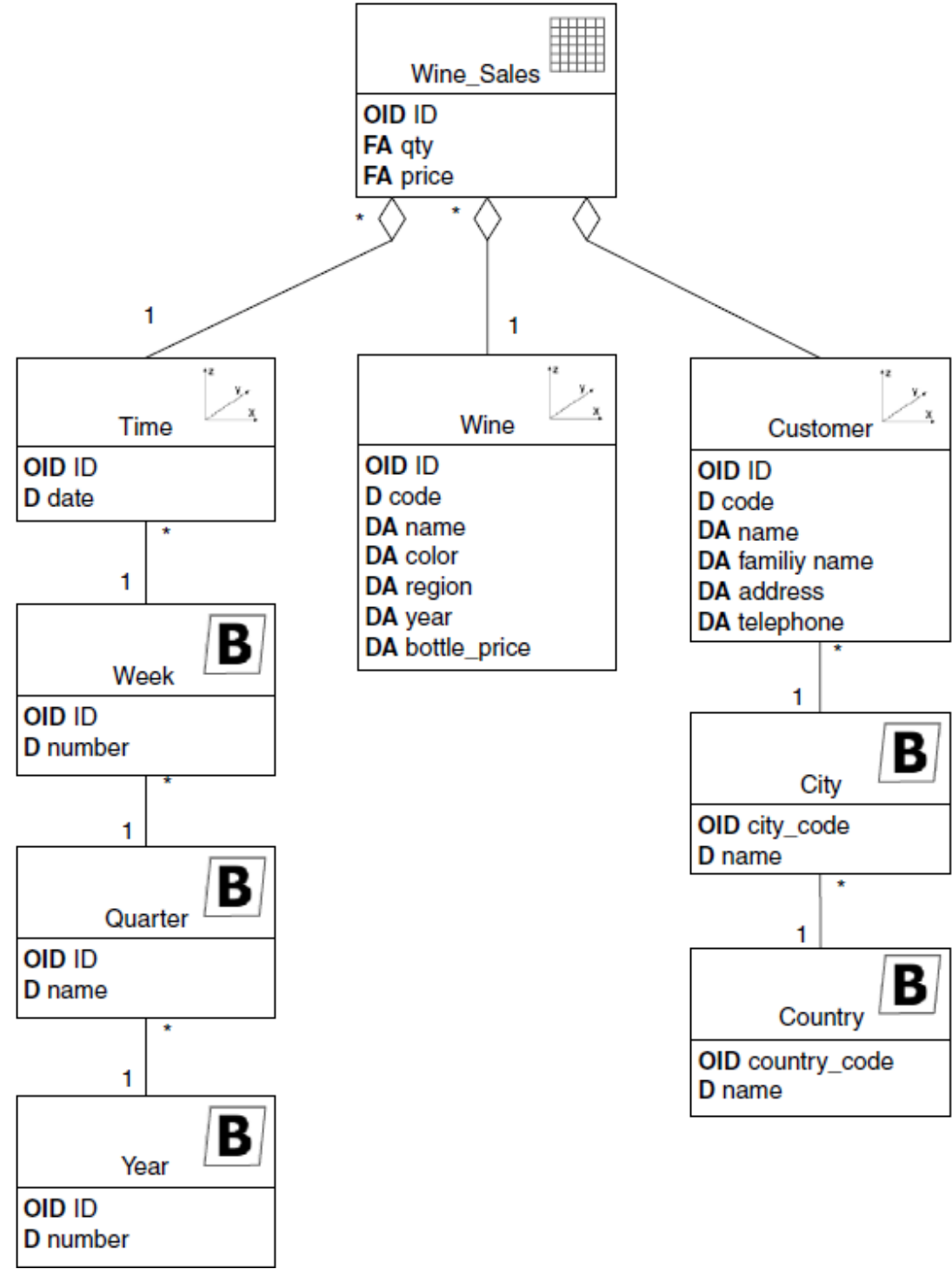


Schéma 60

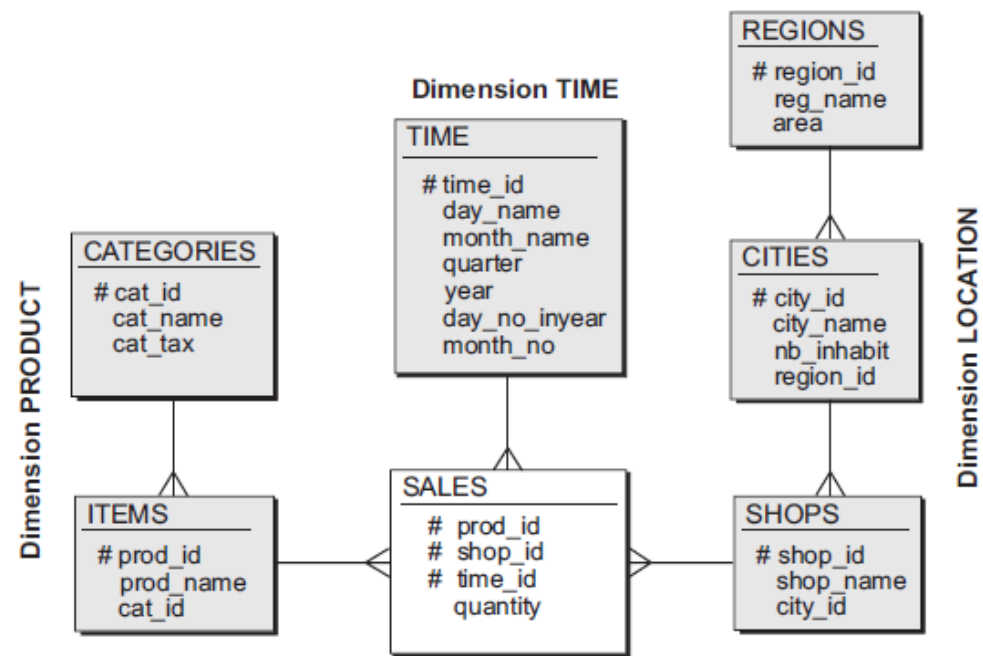
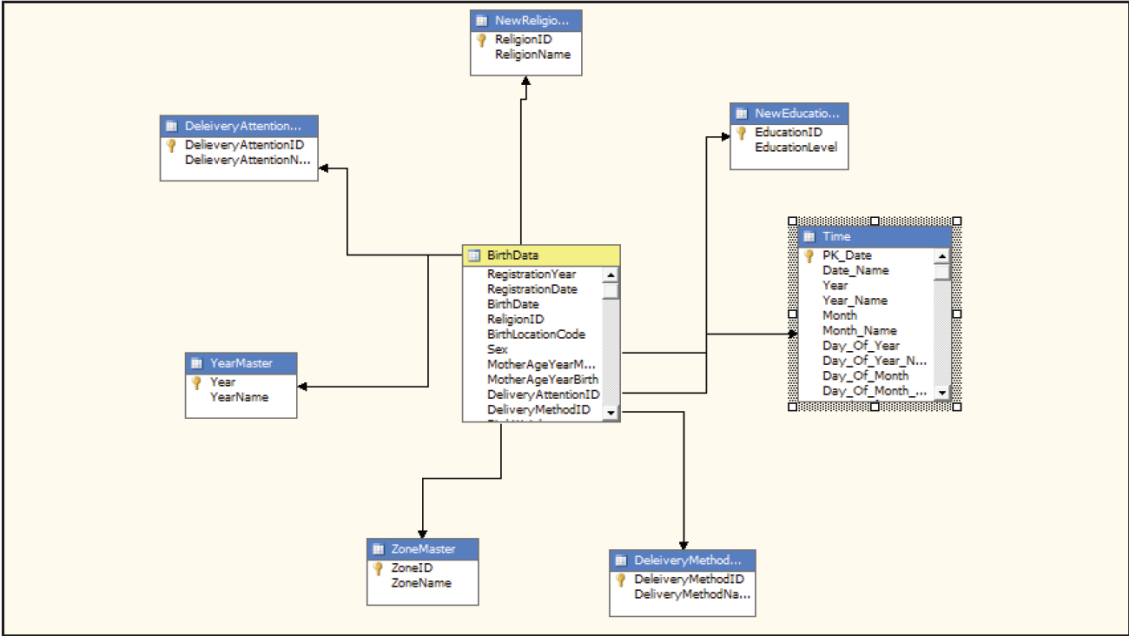


Schéma 61



## Schéma 62

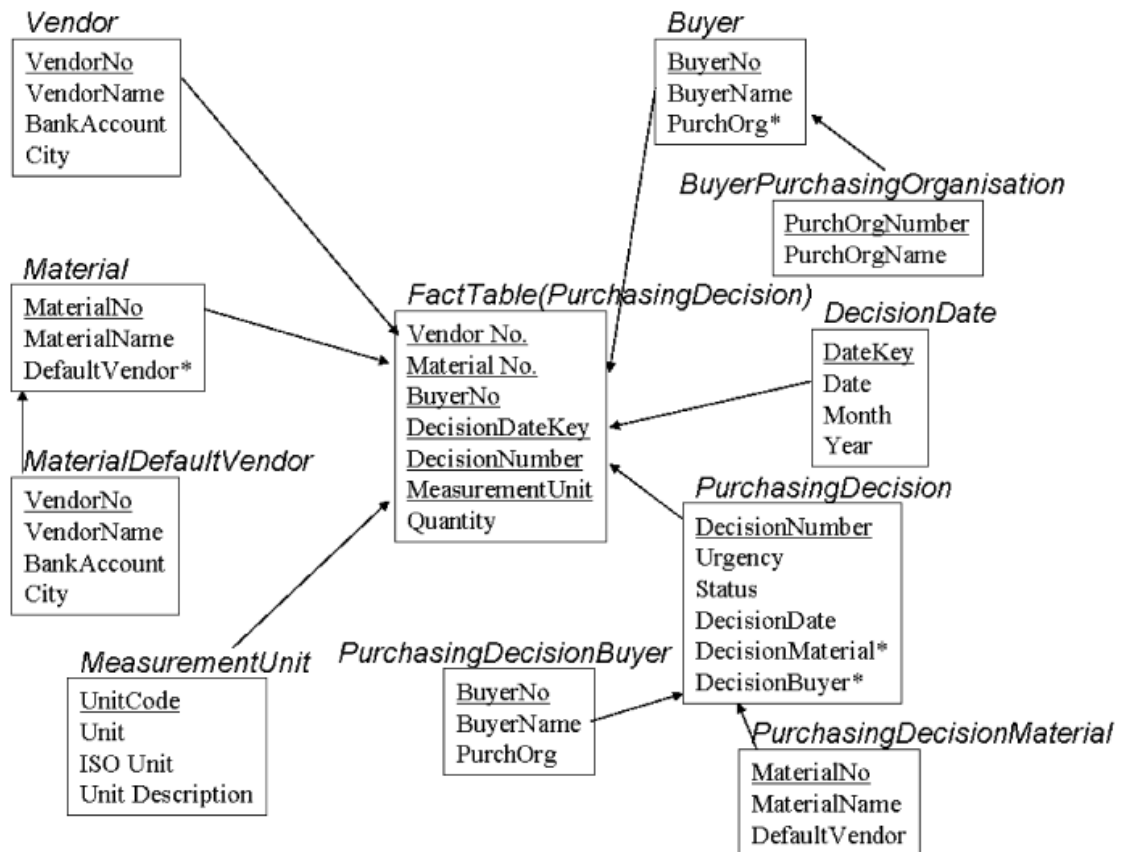


Schéma 63

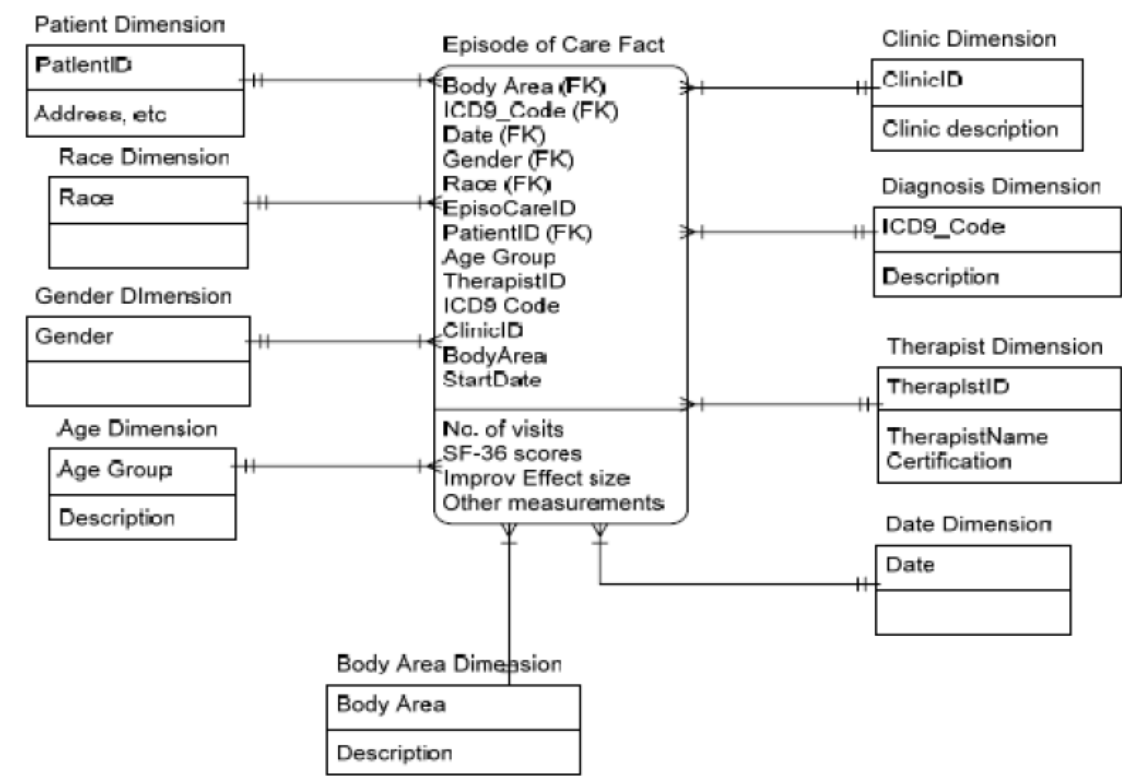
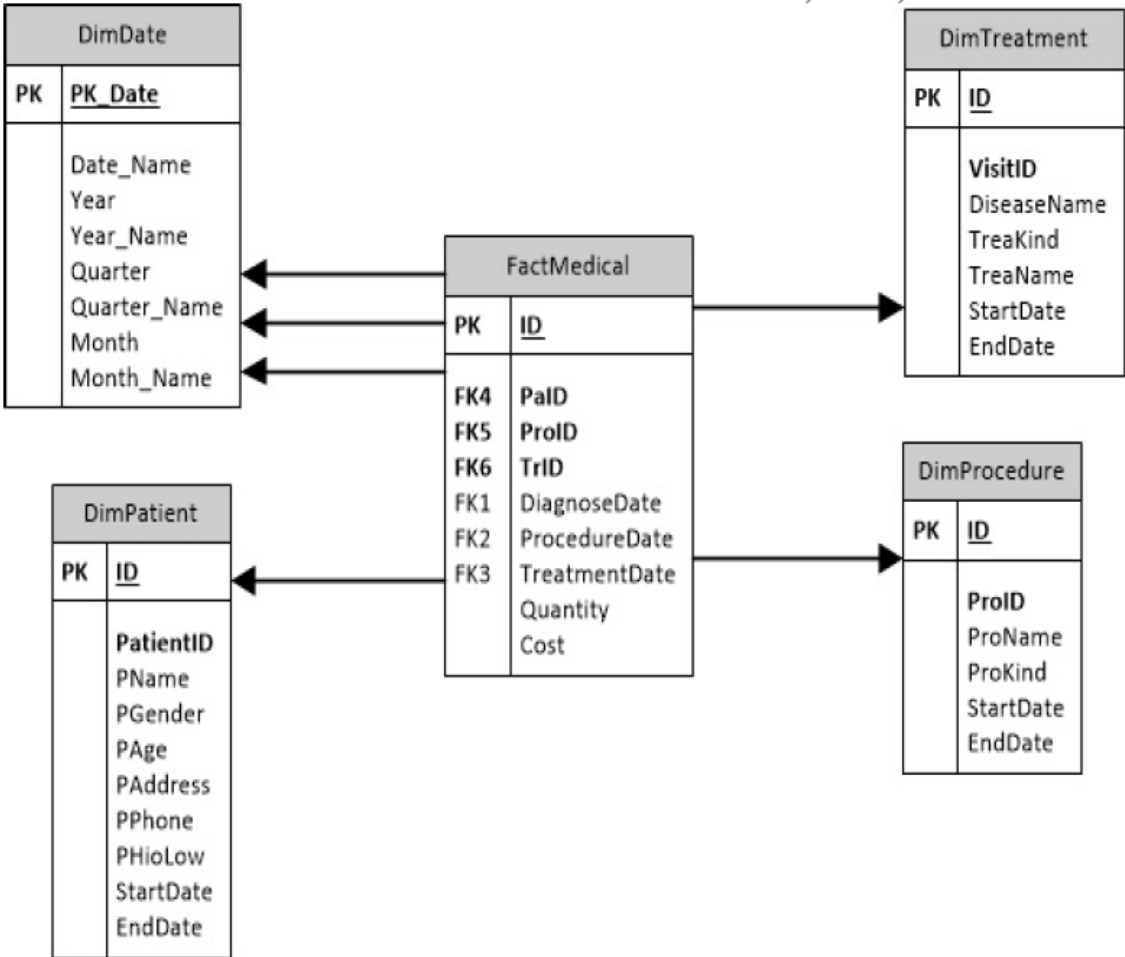
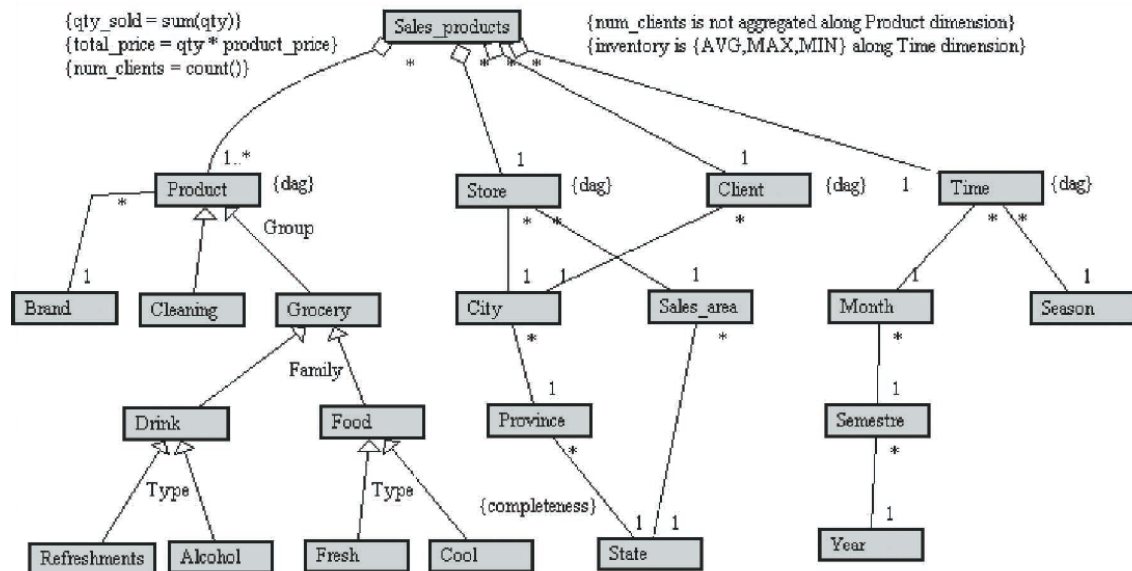


Schéma 64





## Schéma 65



## Schéma 66

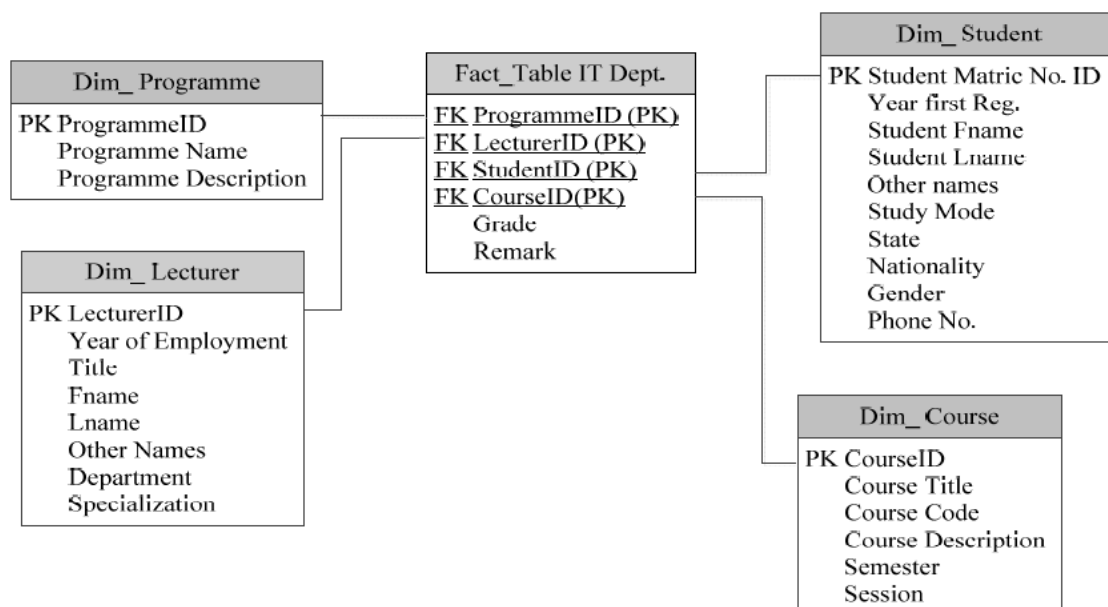


Schéma 67

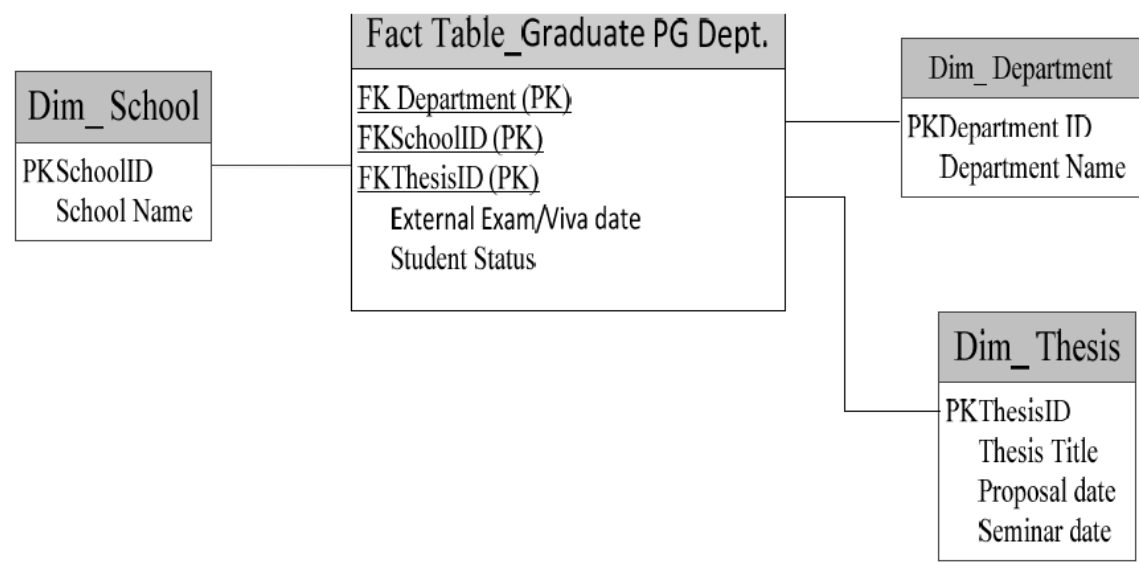


Schéma 68

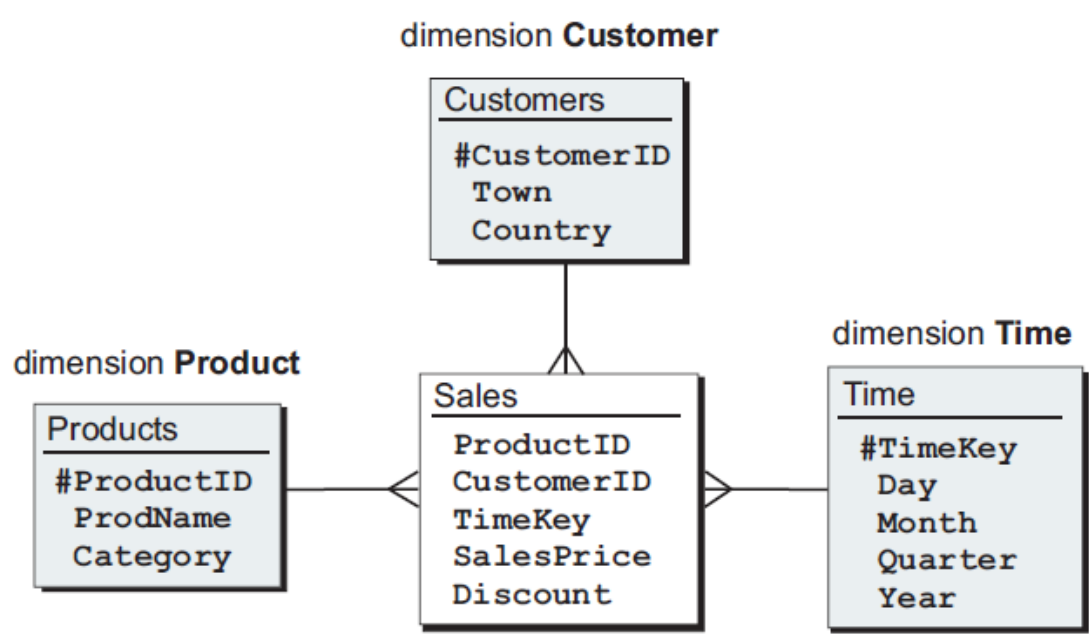


Schéma 69

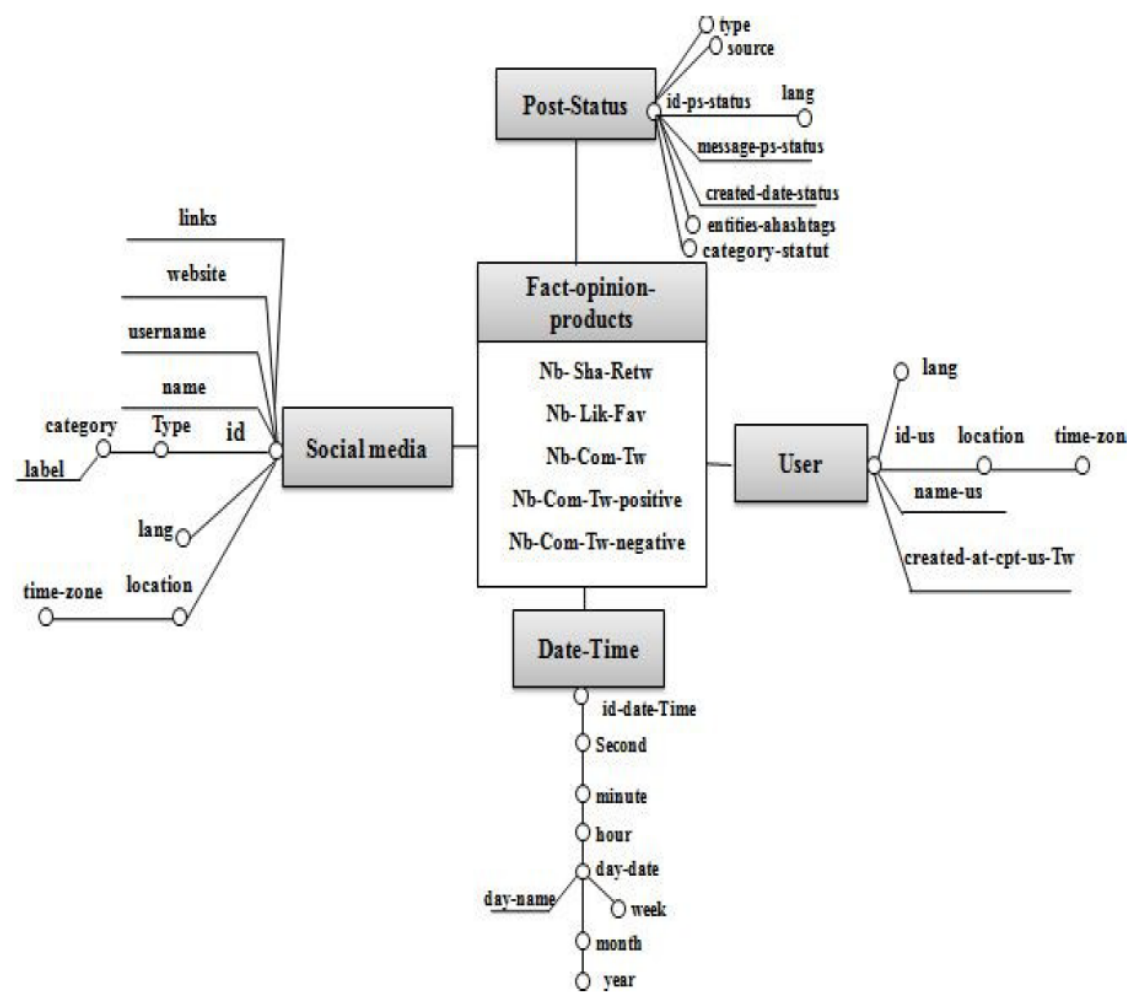


Schéma 70

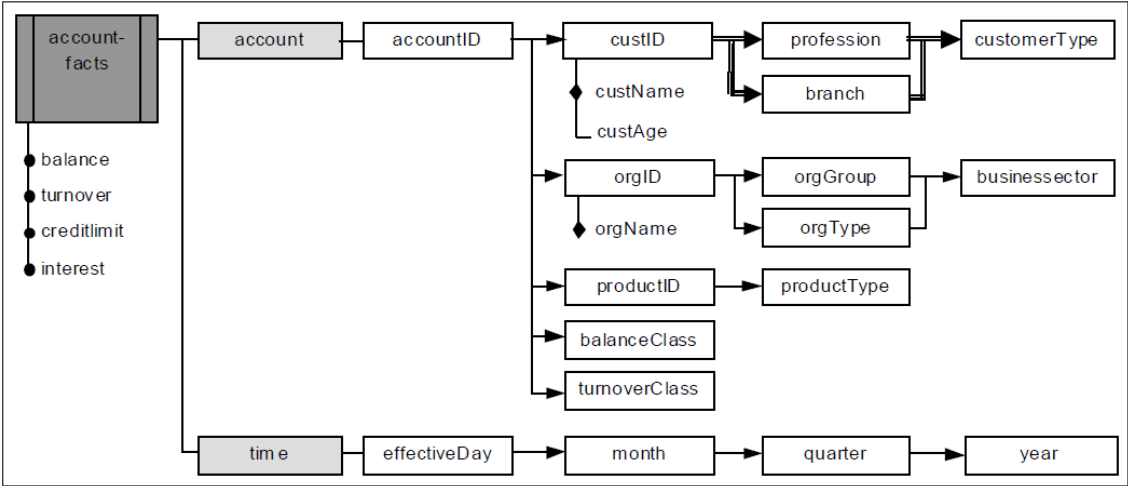


Schéma 71

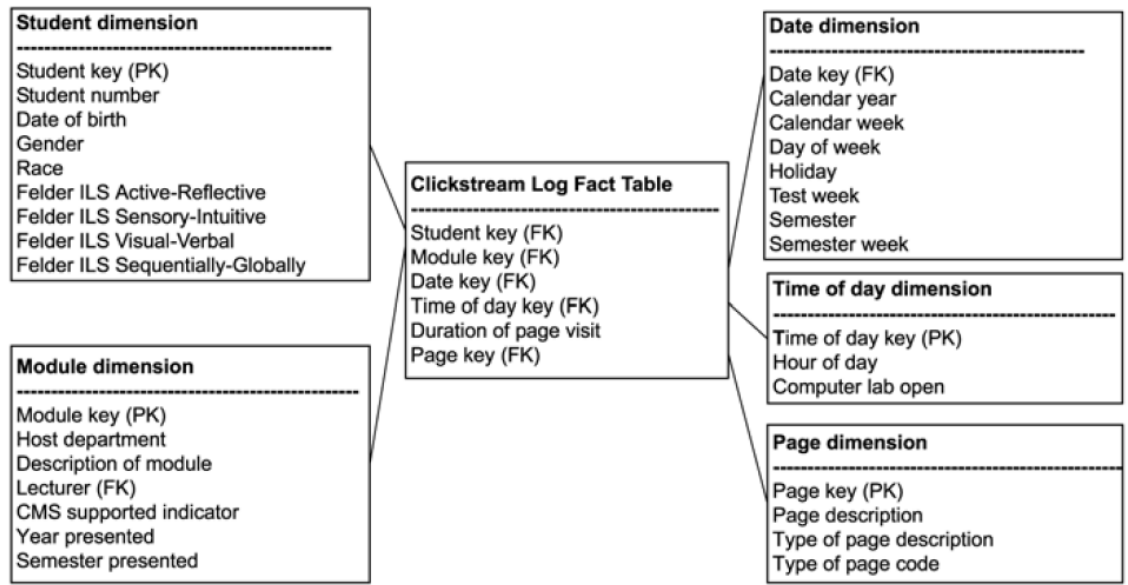


Schéma 72

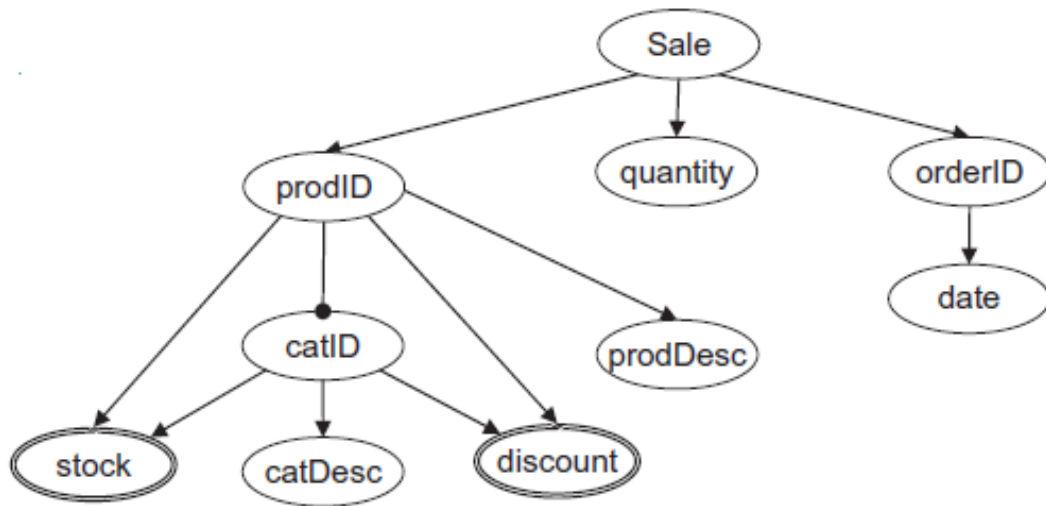
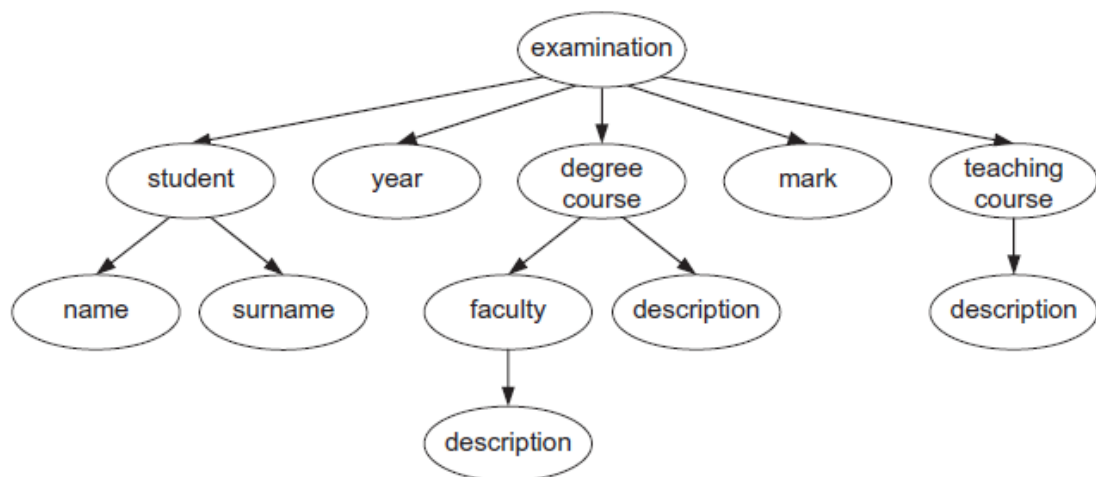


Schéma 73



## Schéma 74

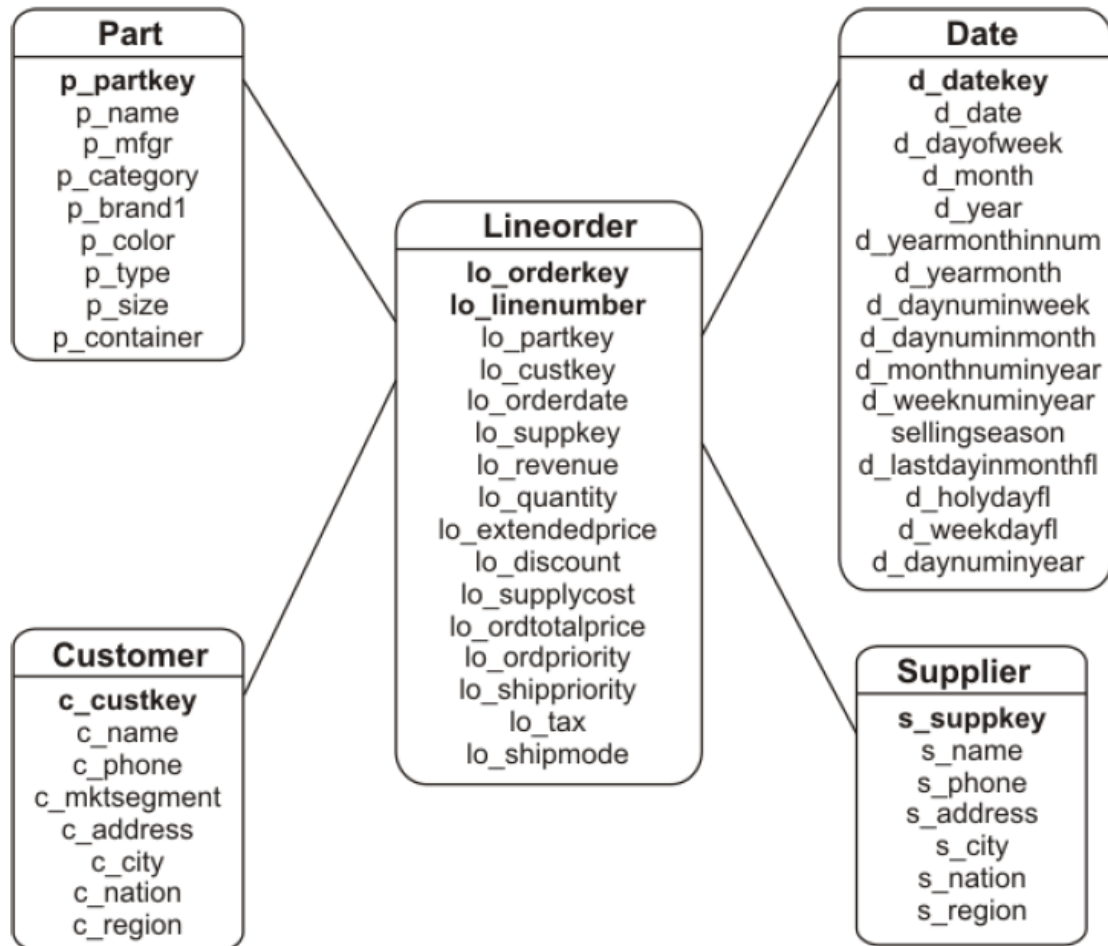


Schéma 75

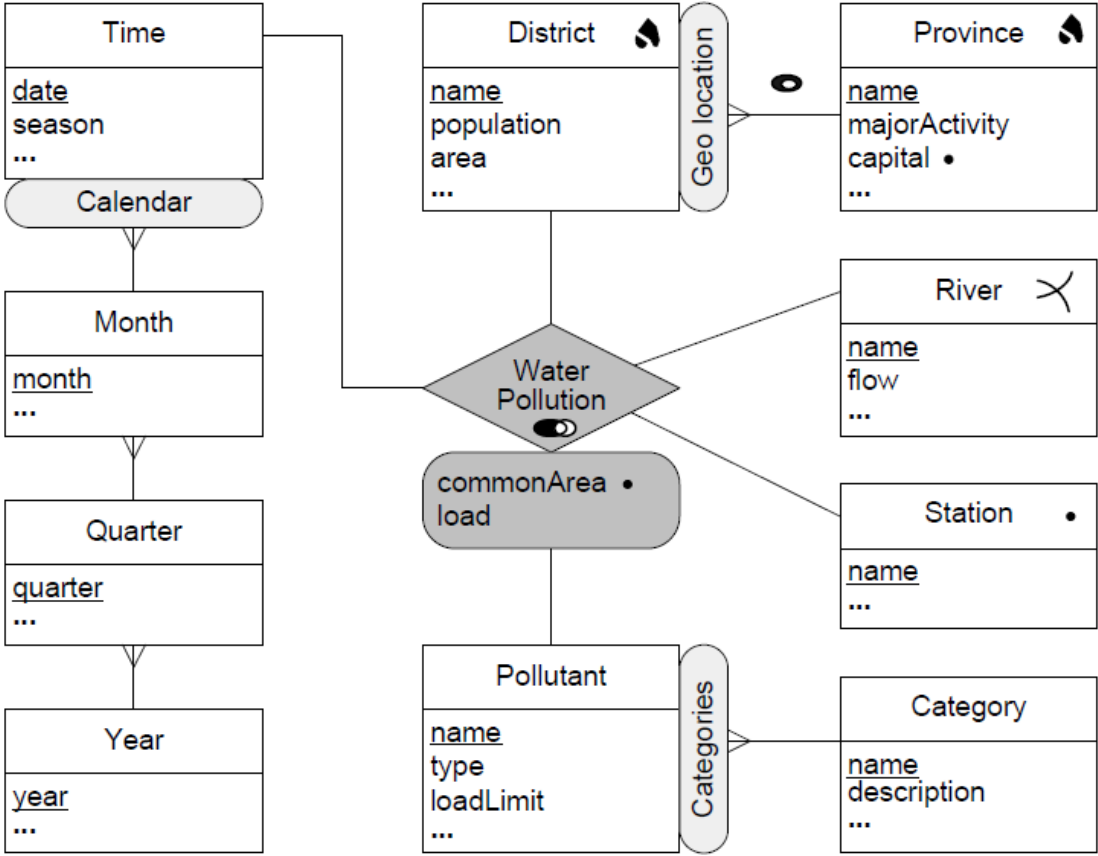


Schéma 76

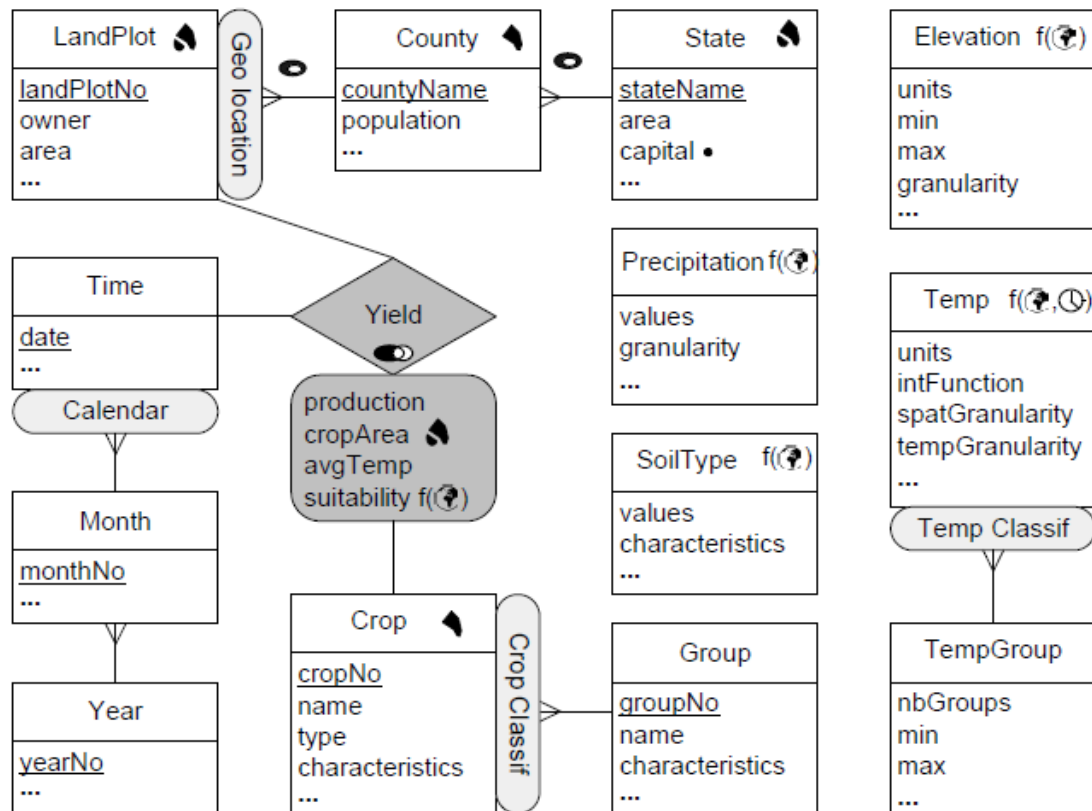




Schéma 77

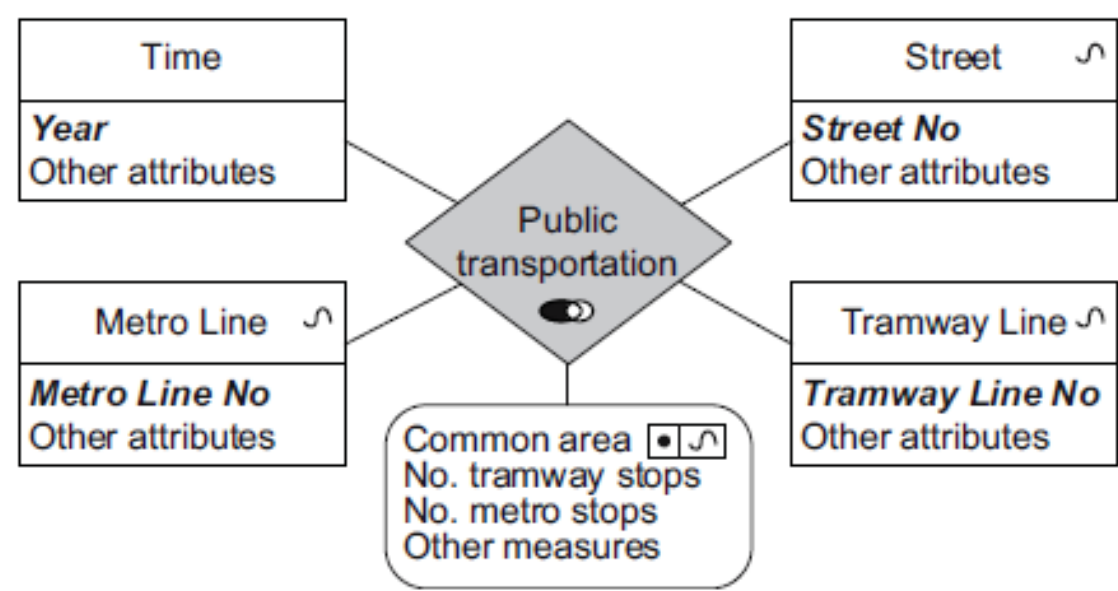
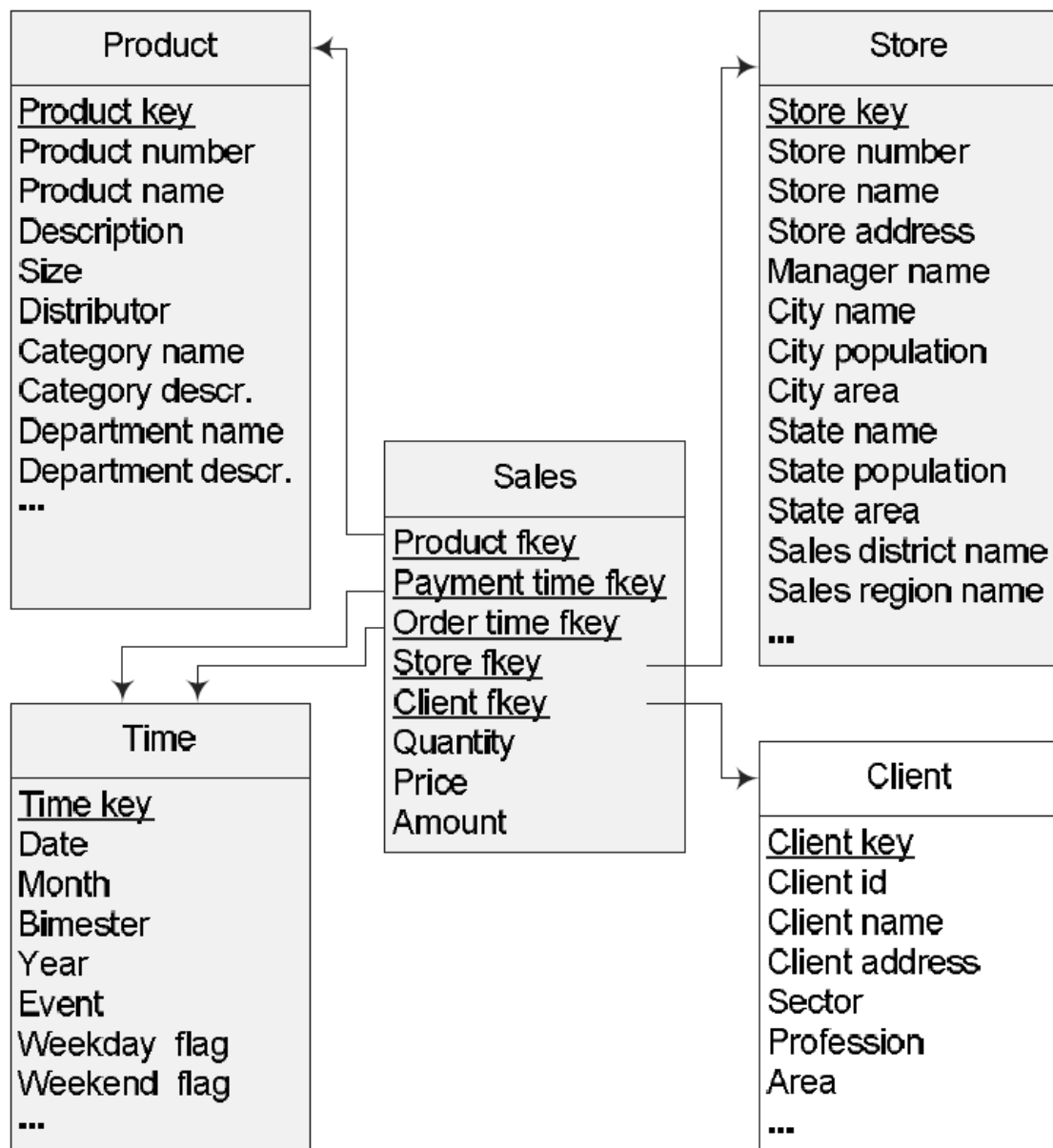


Schéma 78



### Schéma 79

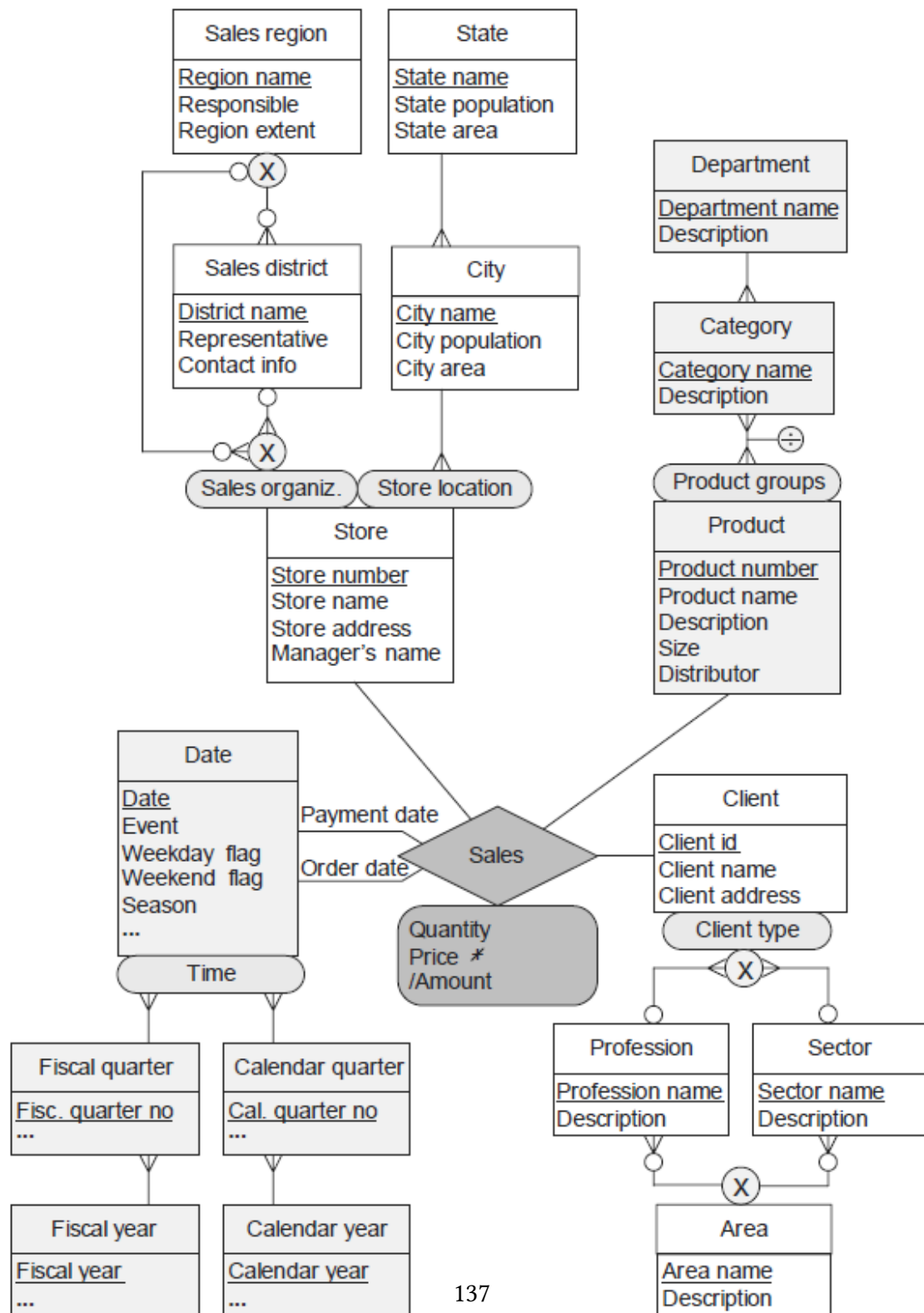
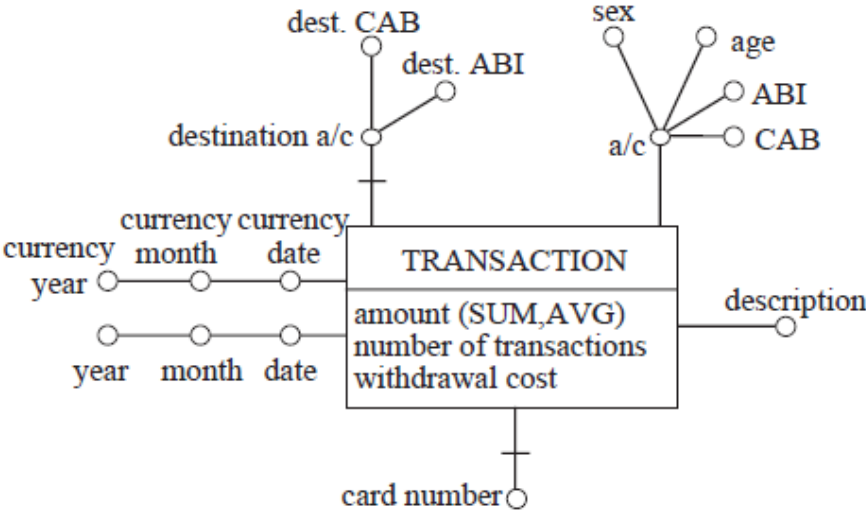


Schéma 80



## Schéma 81

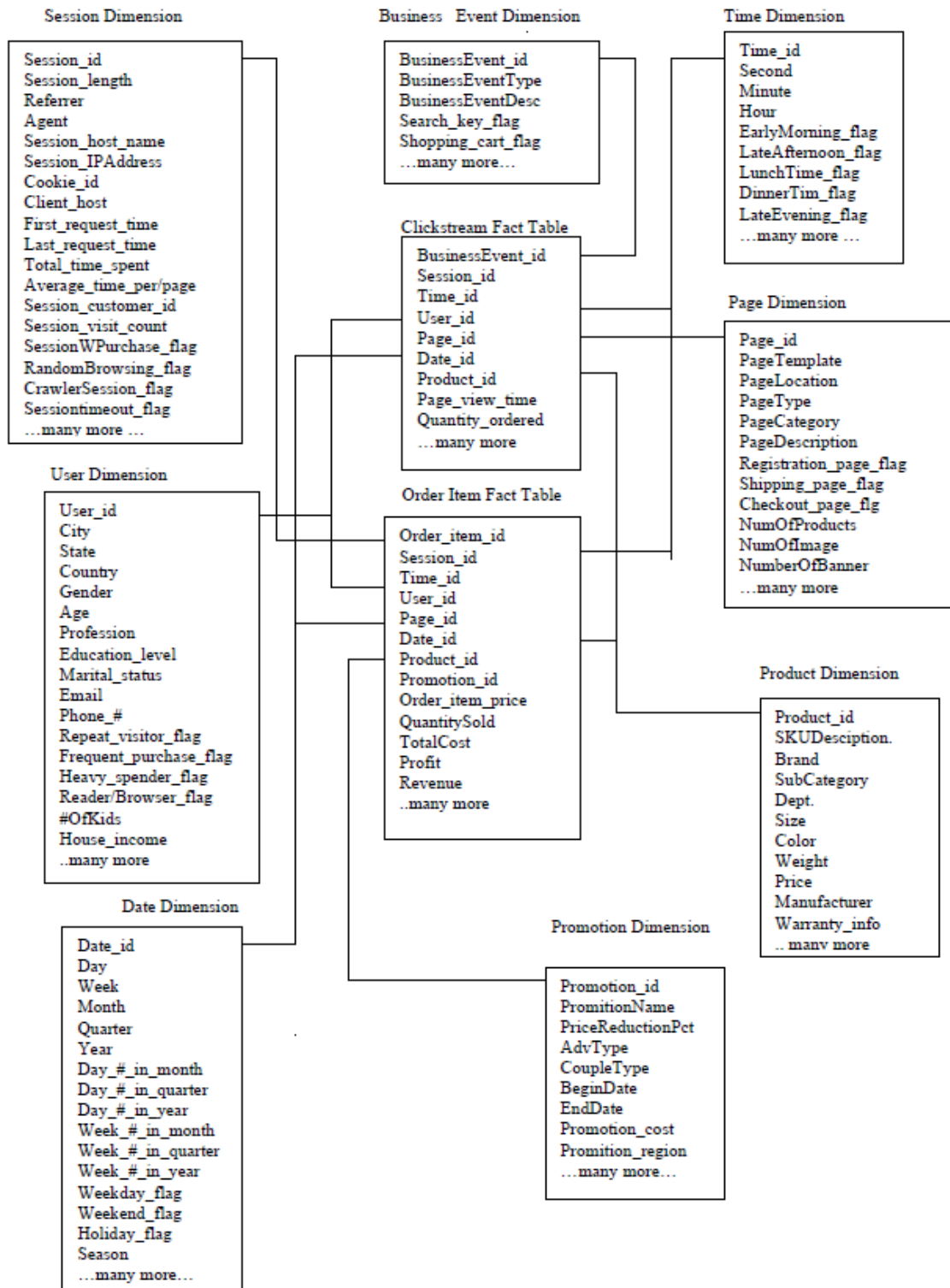


Schéma 82

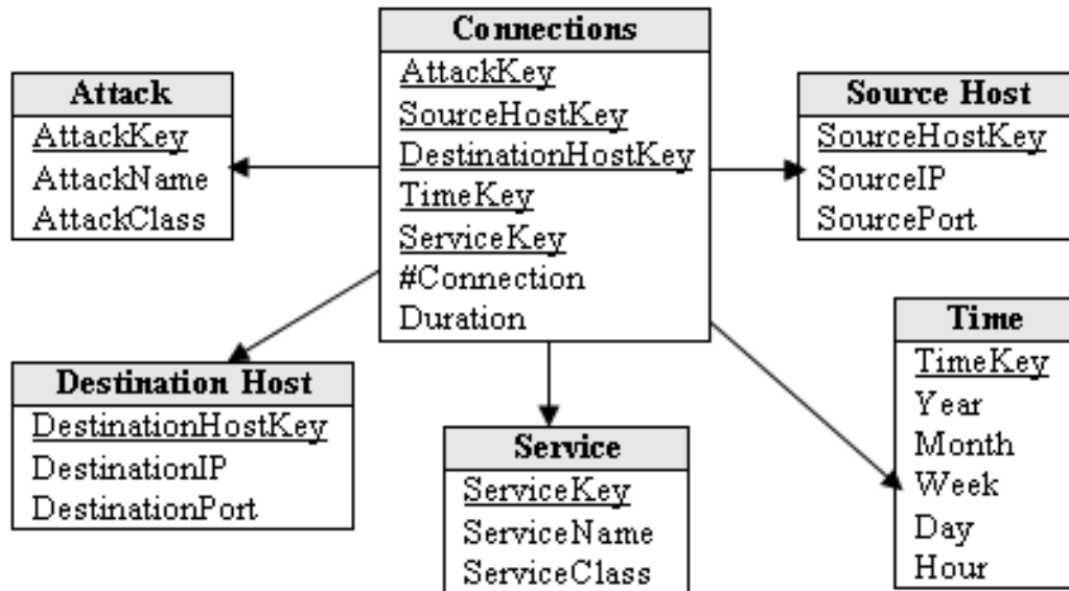


Schéma 83

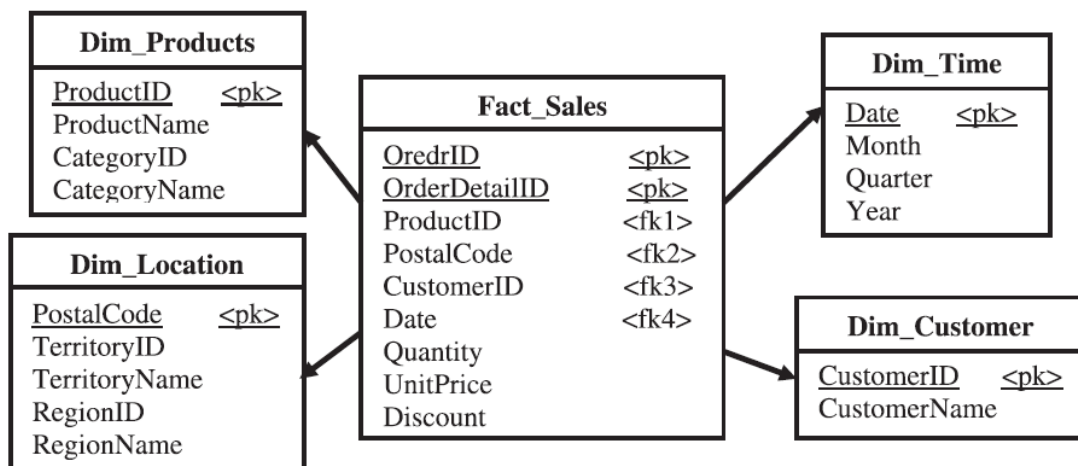


Schéma 84

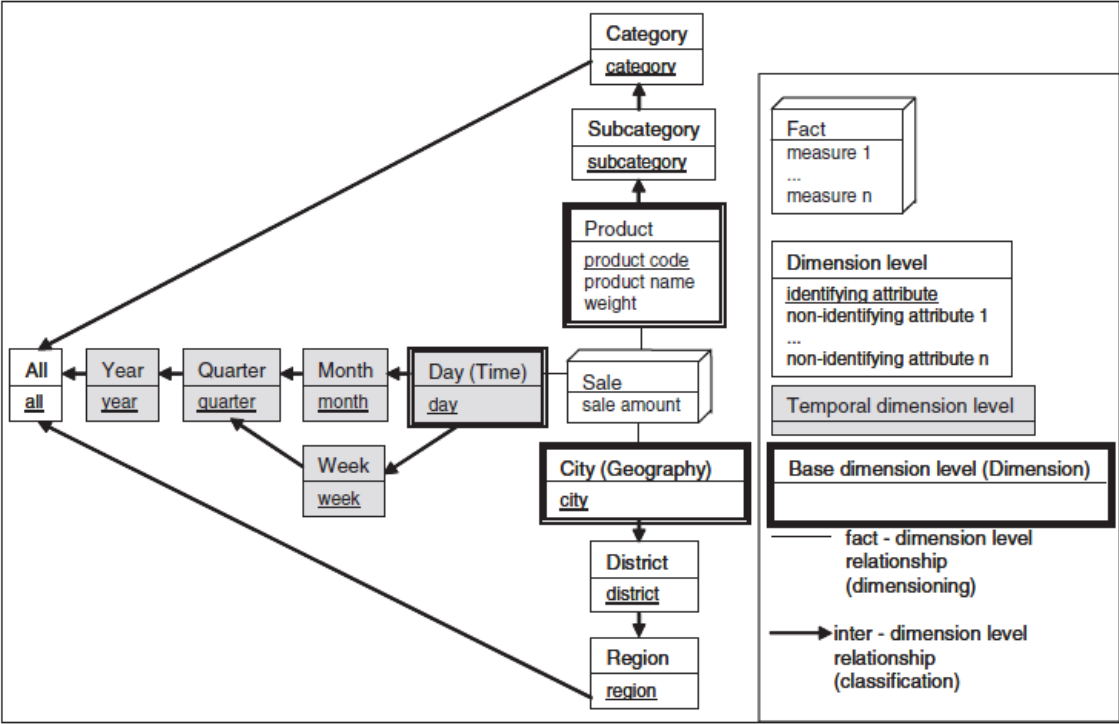


Schéma 85

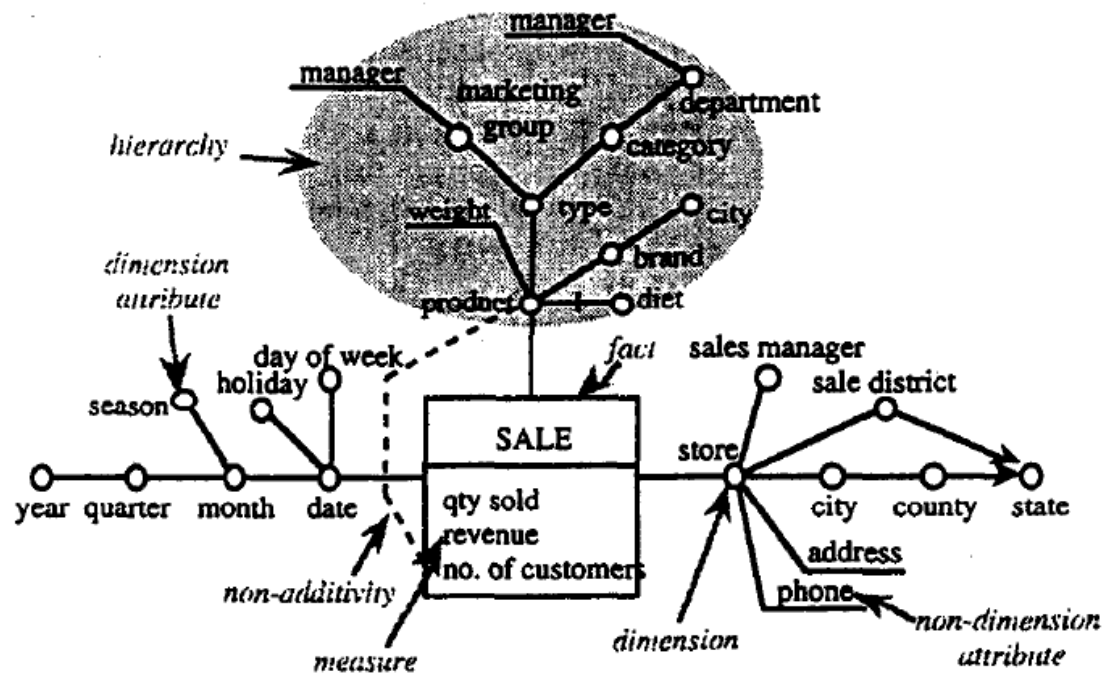




Schéma 86

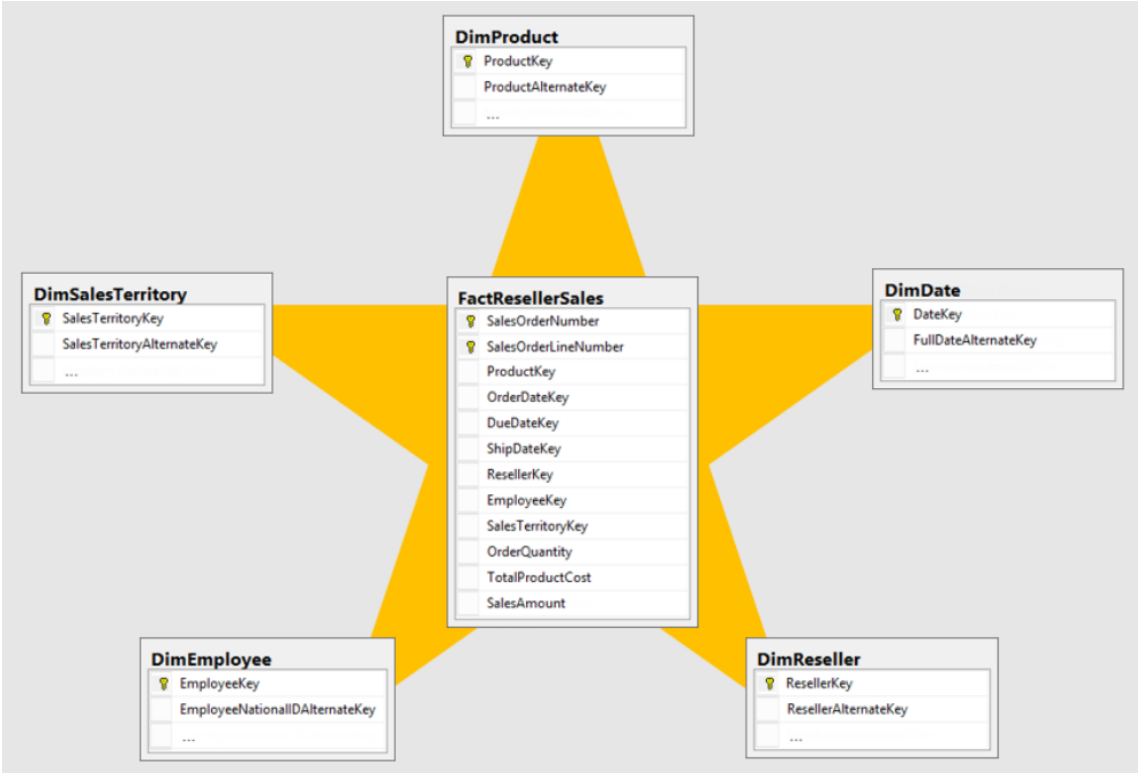
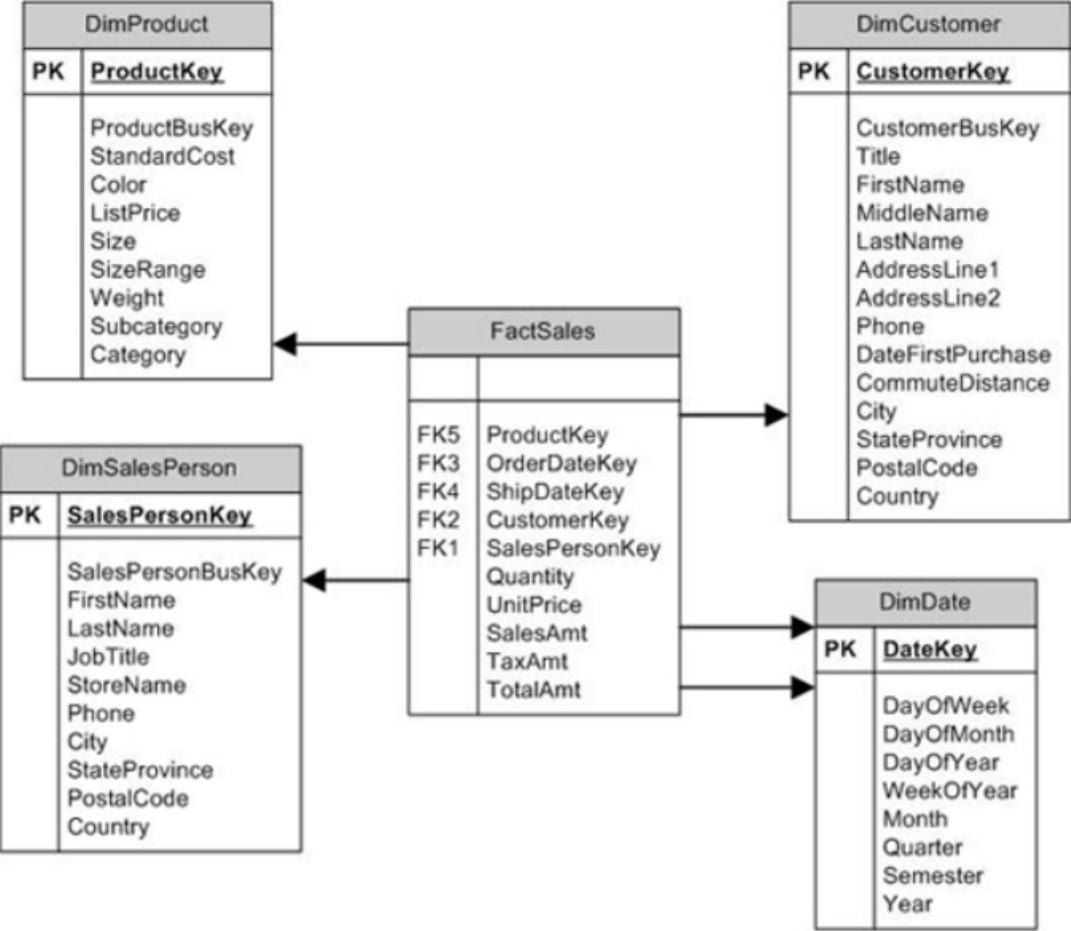


Schéma 87



## Schéma 88

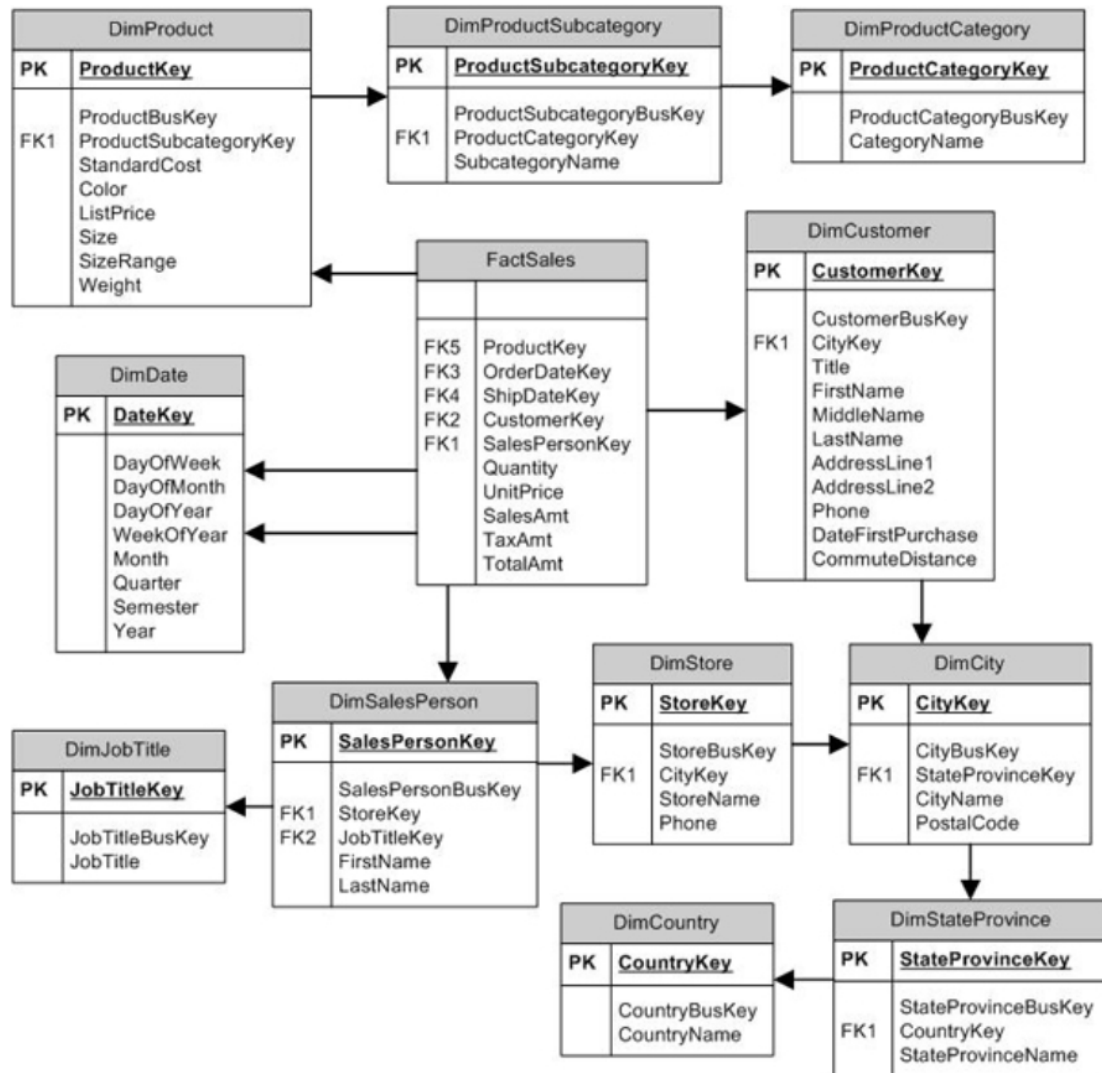


Schéma 89



Schéma 90

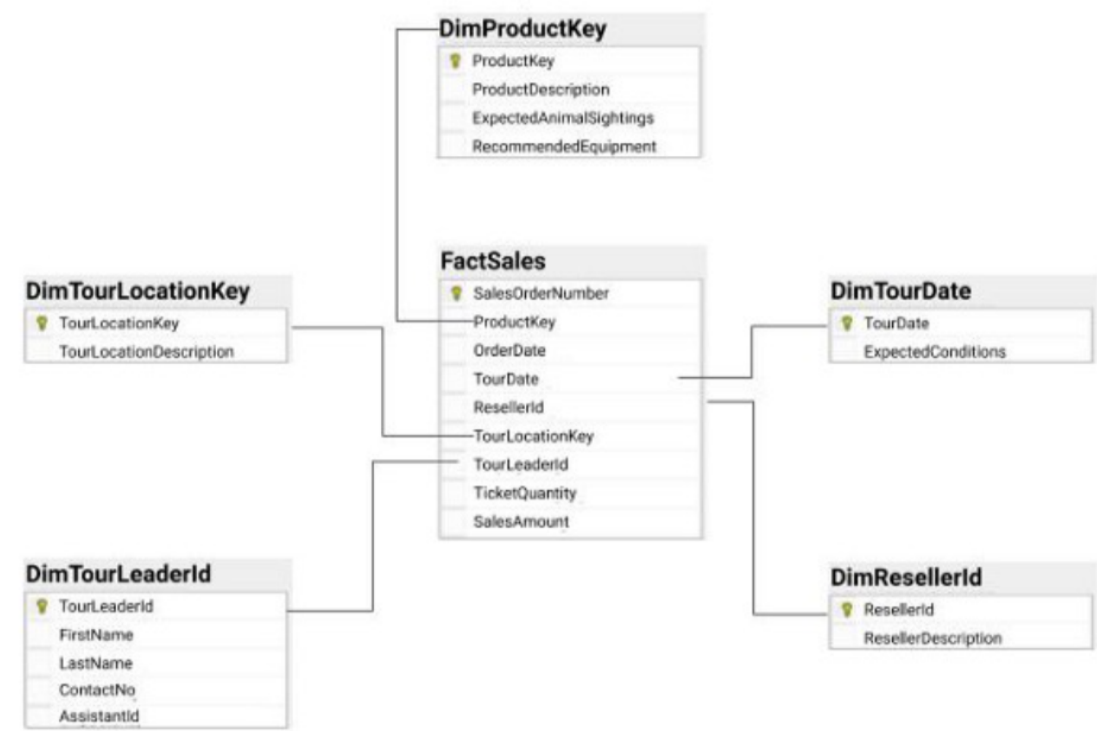
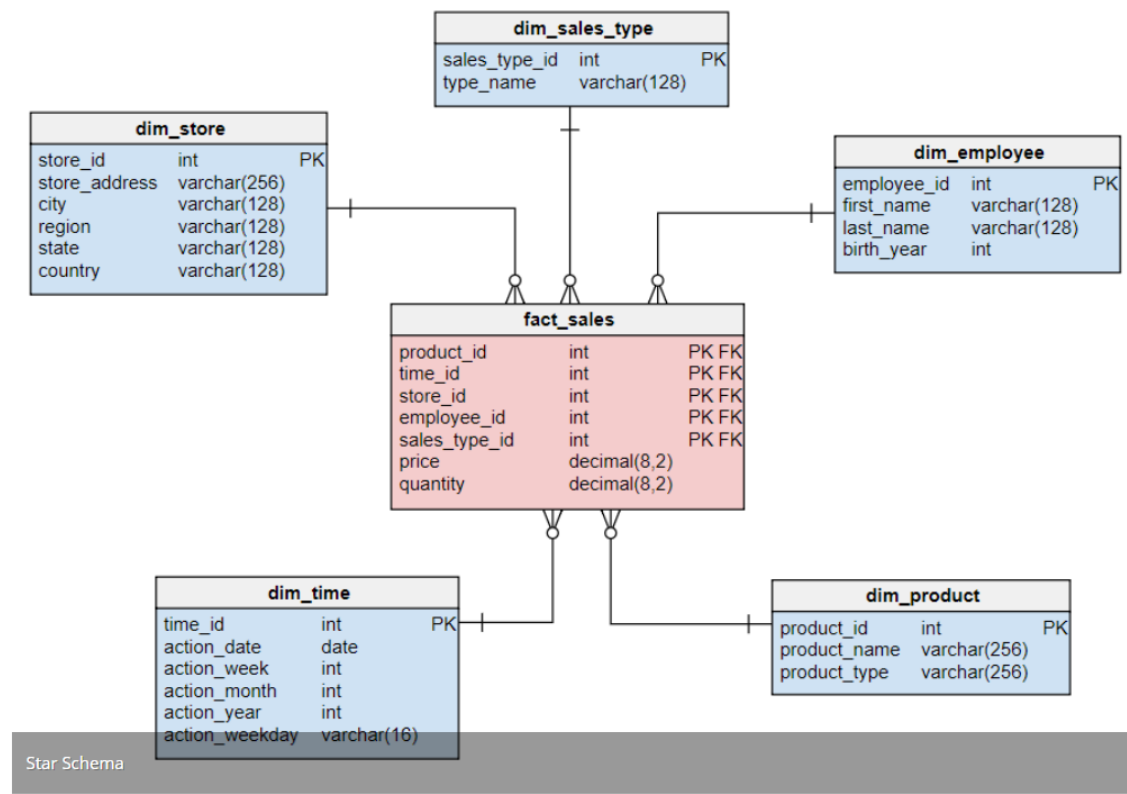


Schéma 91



# Schéma 92

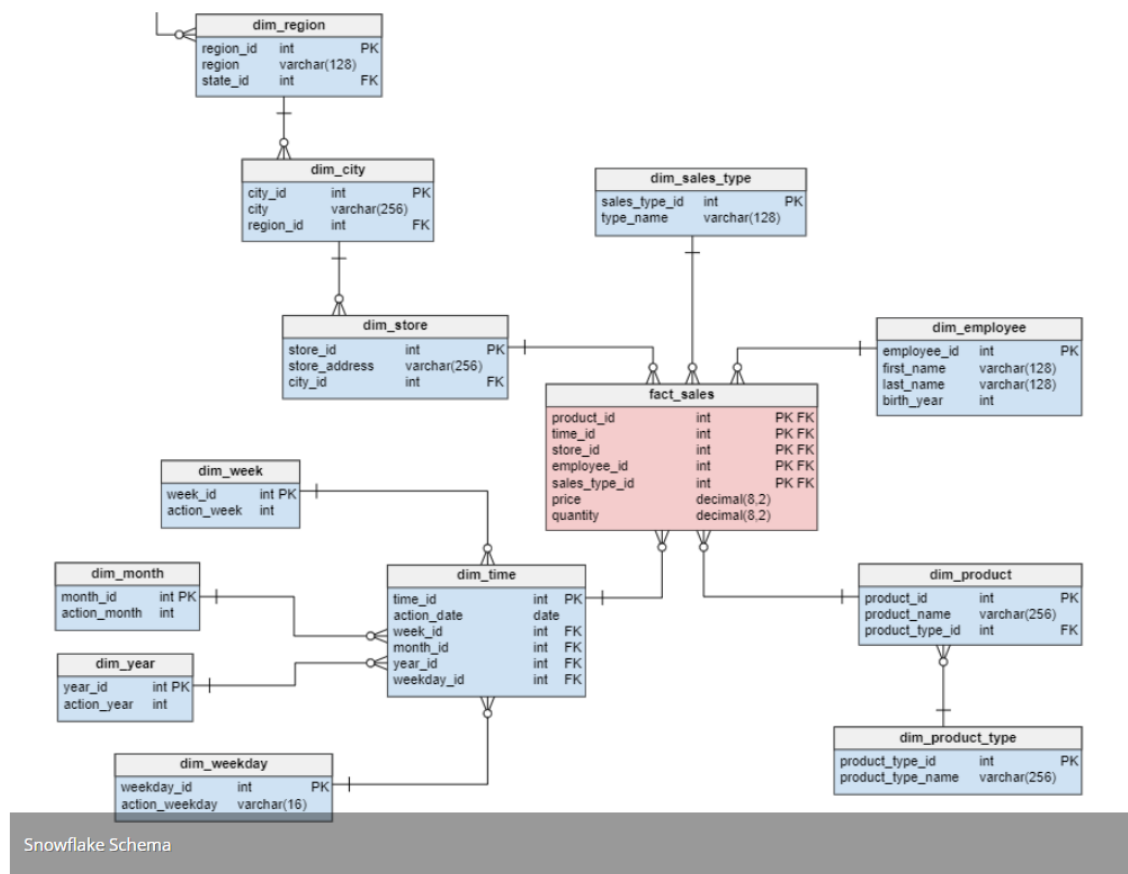
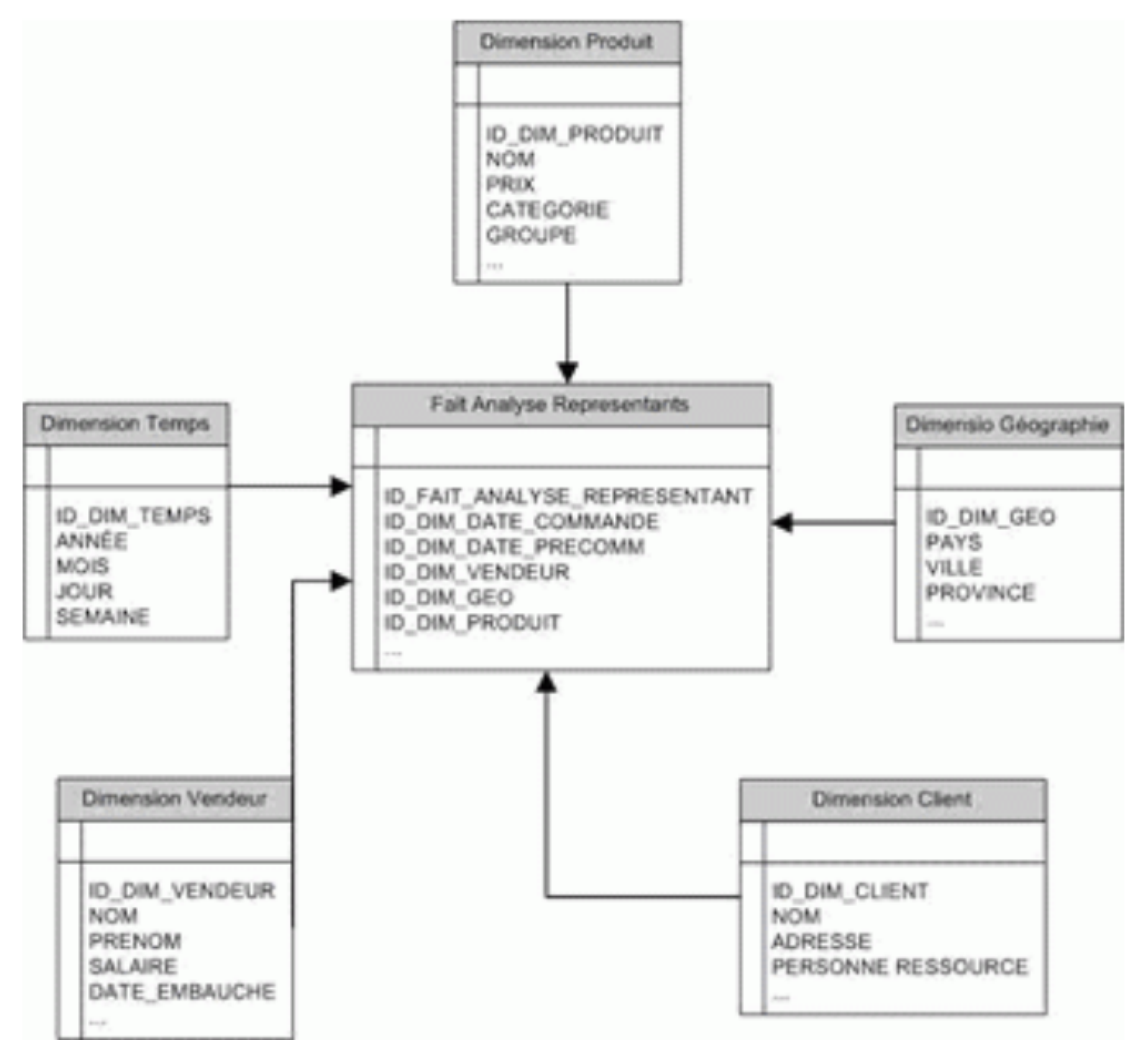


Schéma 93



*Schéma en étoile*



Schéma 94

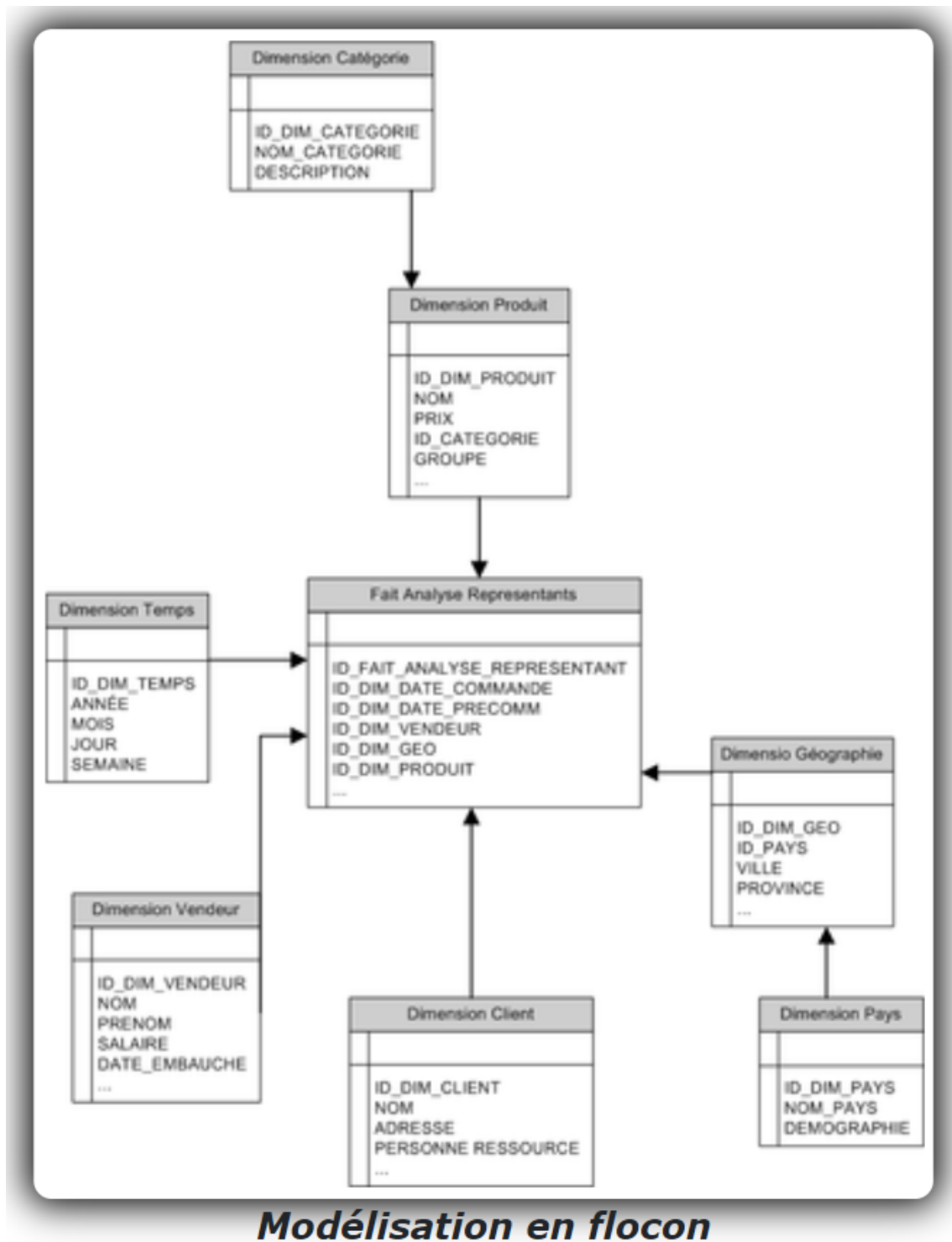
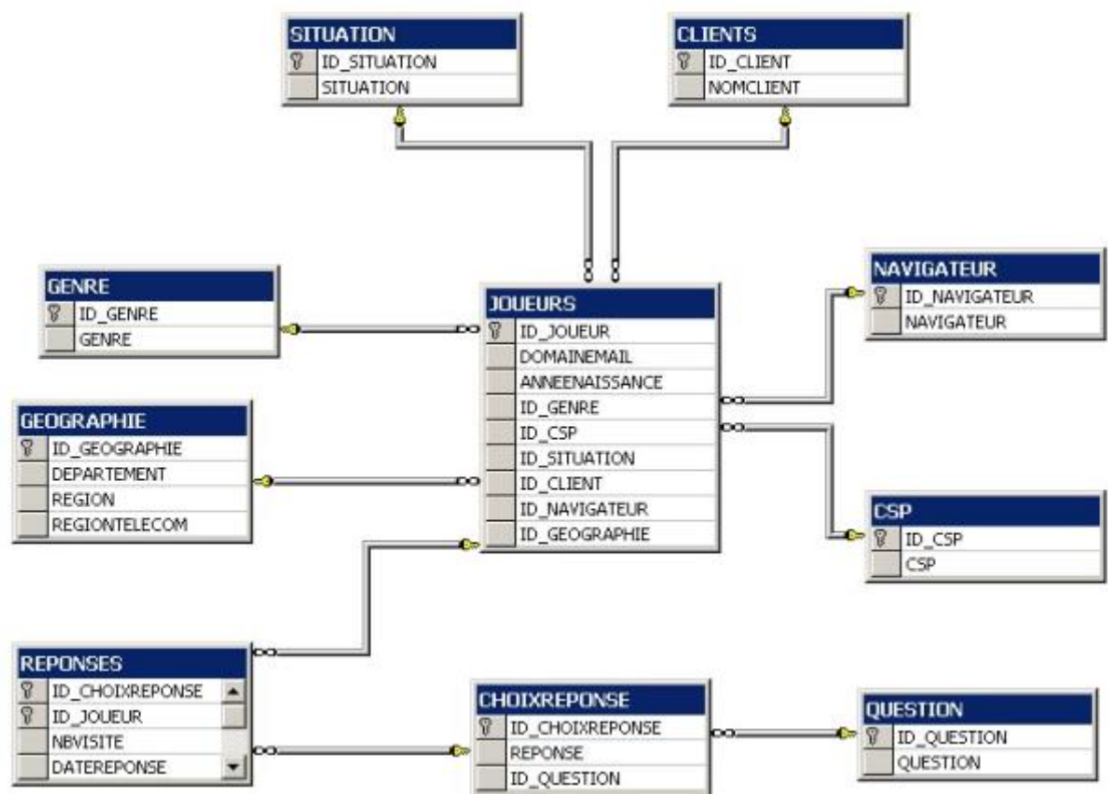


Schéma 95



## Schéma 96

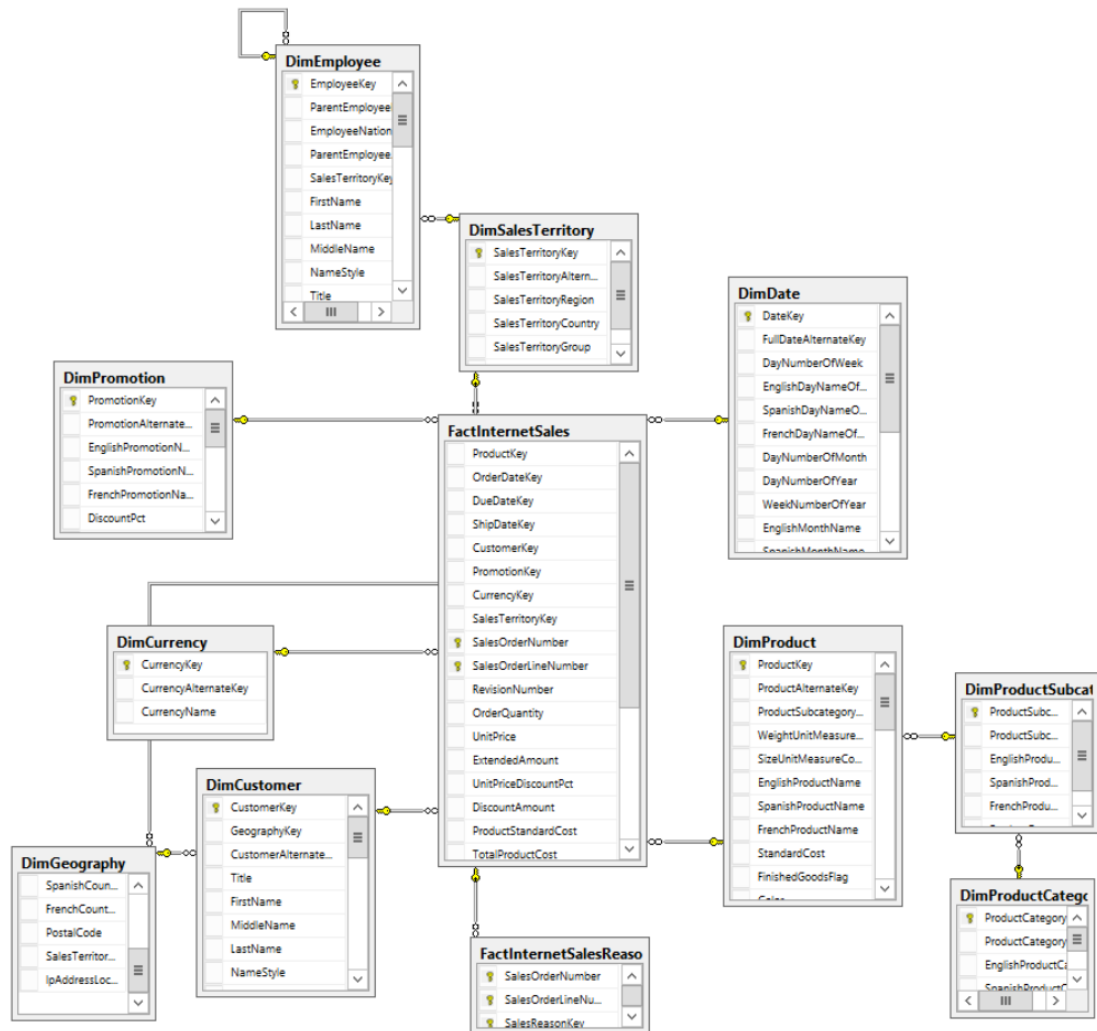


Schéma 97

# Dimensional Modeling

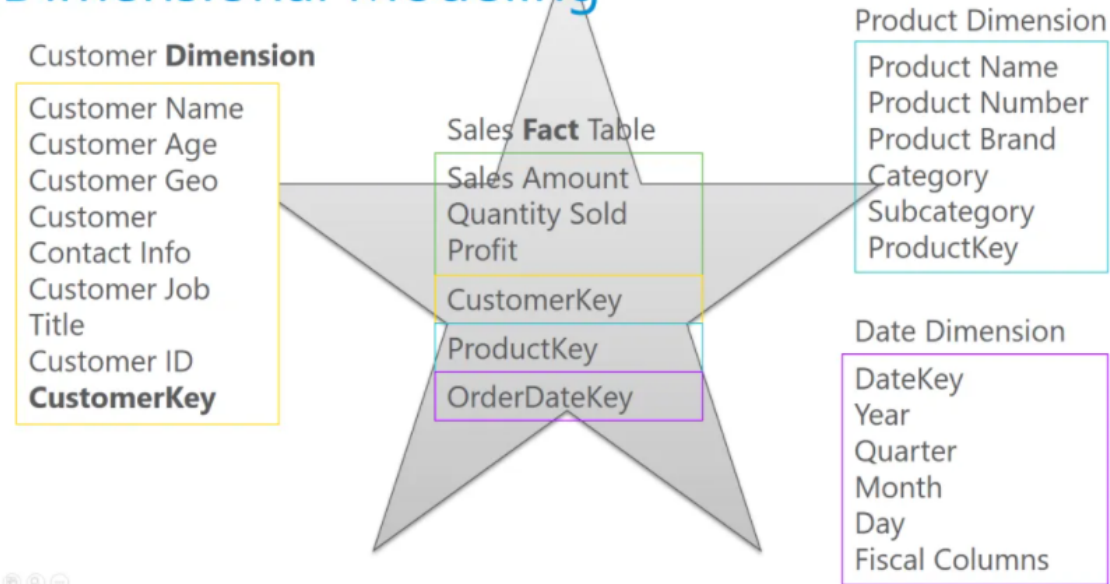


Schéma 98

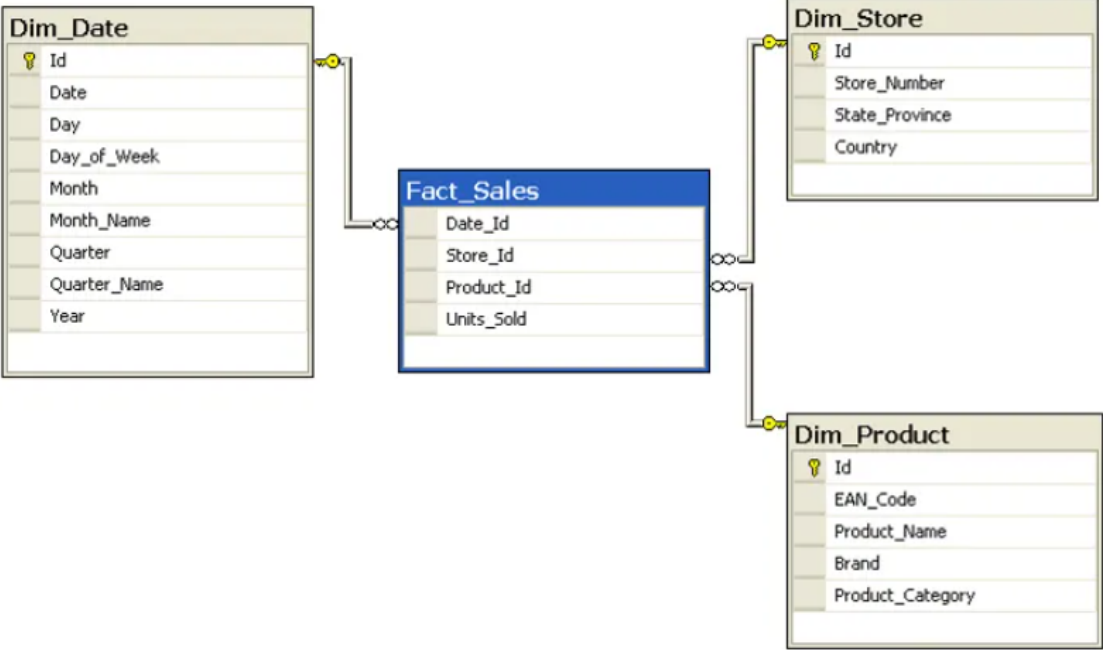


Schéma 99

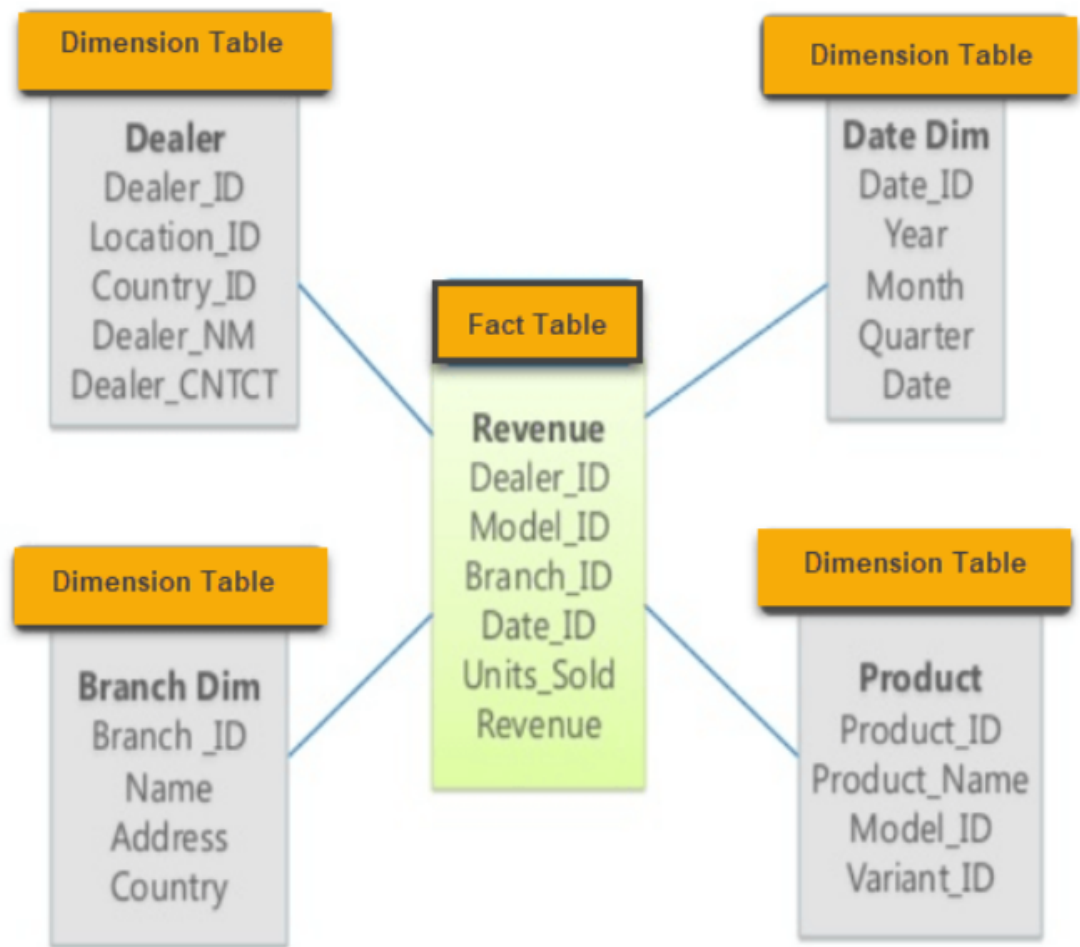
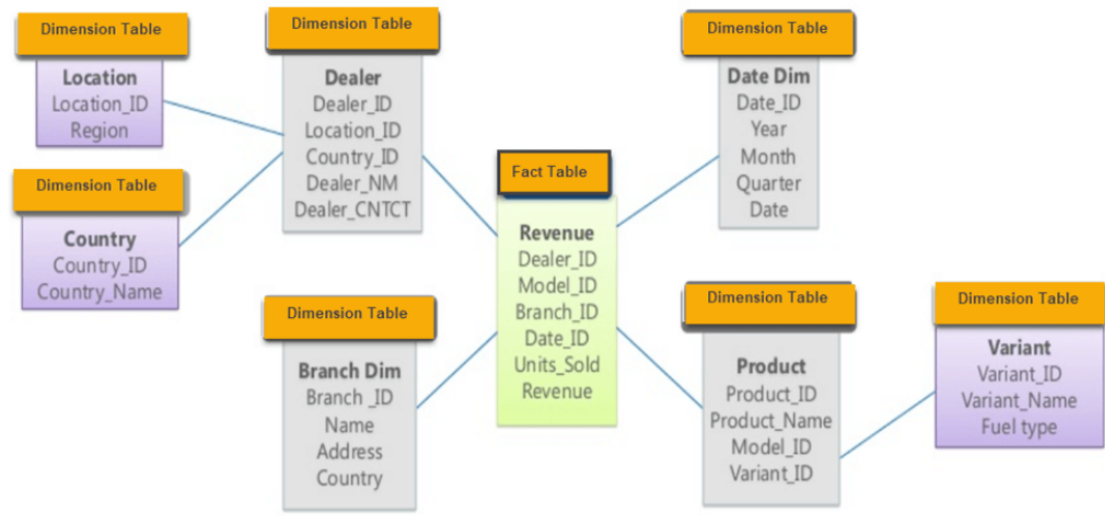


Schéma 100



## Références

- [1] R. Kimball and M. Ross, *The Data Warehouse Toolkit - Third Edition*. Wiley, 2013.
- [2] A. Vaisman and E. Zimányi, *Data warehouse systems*. Springer, 2014.
- [3] M. Golfarelli and S. Rizzi, *Data warehouse design : Modern principles and methodologies*. McGraw-Hill, Inc., 2009.
- [4] B. List, R. M. Bruckner, K. Machaczek, and J. Schiefer, “A comparison of data warehouse development methodologies case study of the process warehouse,” in *International Conference on Database and Expert Systems Applications*. Springer, 2002, pp. 203–215.
- [5] W. H. Inmon, “Building the data warehouse : Getting started,” 2000.
- [6] M. Golfarelli and S. Rizzi, “A methodological framework for data warehouse design,” in *Proceedings of the 1st ACM international workshop on Data warehousing and OLAP*, 1998, pp. 3–9.
- [7] M. Golfarelli, D. Maio, and S. Rizzi, “Conceptual design of data warehouses from e/r schemes,” in *Proceedings of the thirty-first Hawaii international conference on system sciences*, vol. 7. IEEE, 1998, pp. 334–343.
- [8] M. Golfarelli and S. Rizzi, “Designing the data warehouse : Key steps and crucial issues,” *Journal of computer science and Information Management*, vol. 2, no. 3, pp. 88–100, 1999.
- [9] D. L. Moody and M. A. Kortink, “From enterprise models to dimensional models : a methodology for data warehouse and data mart design.” in *DMDW*, 2000, p. 5.
- [10] S. Luján-Mora and J. Trujillo, “A data warehouse engineering process,” in *International Conference on Advances in Information Systems*. Springer, 2004, pp. 14–23.
- [11] B. Hüsemann, J. Lechtenbörger, and G. Vossen, *Conceptual data warehouse design*. Universität Münster. Angewandte Mathematik und Informatik, 2000.



- [12] J. Trujillo, M. Palomar, J. Gomez, and I.-Y. Song, "Designing data warehouses with oo conceptual models," *Computer*, vol. 34, no. 12, pp. 66–75, 2001.
- [13] E. Malinowski and E. Zimányi, "Conceptual modeling for data warehouse and olap applications," in *Encyclopedia of Data Warehousing and Mining, Second Edition*. IGI Global, 2009, pp. 293–300.
- [14] Les design patterns : une programmation plus rapide et plus sûre.  
[Https ://www.ionos.fr/digitalguide/sites-internet/developpement-web/les-design-patterns-quest-ce-que-cest/](https://www.ionos.fr/digitalguide/sites-internet/developpement-web/les-design-patterns-quest-ce-que-cest/).
- [15] Christopher alexander. [Https ://fr.wikipedia.org/wiki/Christopher-Alexander](https://fr.wikipedia.org/wiki/Christopher-Alexander).
- [16] T. Mikkonen, "Formalizing design patterns," in *Proceedings of the 20th international conference on Software engineering*. IEEE, 1998, pp. 115–124.
- [17] S. Rahman. The 3 types of design patterns all developers should know.  
[Https ://www.freecodecamp.org/news/the-basic-design-patterns-all-developers-need-to-know/](https://www.freecodecamp.org/news/the-basic-design-patterns-all-developers-need-to-know/).
- [18] J. Heer and M. Agrawala, "Software design patterns for information visualization," *IEEE transactions on visualization and computer graphics*, vol. 12, no. 5, pp. 853–860, 2006.
- [19] J. Dong, Y. Sun, and Y. Zhao, "Design pattern detection by template matching," in *Proceedings of the 2008 ACM symposium on Applied computing*, 2008, pp. 765–769.
- [20] A. Fardeay. Qu'est-ce qu'un design pattern ?  
[Https ://www.geeksforgeeks.org/design-patterns-set-1-introduction/](https://www.geeksforgeeks.org/design-patterns-set-1-introduction/).
- [21] A. Saha and T. Praharaj. Design patterns. [Https ://www.geeksforgeeks.org/design-patterns-set-1-introduction/](https://www.geeksforgeeks.org/design-patterns-set-1-introduction/).
- [22] K. Gayantha. Client-server architectural pattern.  
[Https ://medium.com/@96kavindugayantha/software-architectural-patterns-17b05bfa3aef](https://medium.com/@96kavindugayantha/software-architectural-patterns-17b05bfa3aef).

- [23] V. Mallawaarachchi. 10 common software architectural patterns in a nutshell. <https://towardsdatascience.com/10-common-software-architectural-patterns-in-a-nutshell-a0b47a1e9013>.
- [24] T. J. Marlowe, C. S. Ku, and J. W. Benham, "Design patterns for database pedagogy : a proposal," *ACM SIGCSE Bulletin*, vol. 37, no. 1, pp. 48–52, 2005.
- [25] R. A. Haraty and G. Stephan, "Relational database design patterns," in *2013 IEEE 16th International Conference on Computational Science and Engineering*. IEEE, 2013, pp. 818–824.
- [26] M. N. Shirazi, H. C. Kuan, and H. Dolatabadi, "Design patterns to enable data portability between clouds' databases," in *2012 12th International Conference on Computational Science and Its Applications*. IEEE, 2012, pp. 117–120.
- [27] B. Design. Qu'est-ce que la rétro-ingénierie? <http://www.bpdesignfrance.com/quest-ce-que-la-retro-ingenierie/>.
- [28] Y. Song, F. Pramudianto, and E. F. Gehringer, "A markup language for building a data warehouse for educational peer-assessment research," in *2016 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2016, pp. 1–5.
- [29] T. Ishaya and M. Folarin, "A service oriented approach to business intelligence in telecoms industry," *Telematics and Informatics*, vol. 29, no. 3, pp. 273–285, 2012.
- [30] T. F. Di, E. Lefons, and F. Tangorra, "Academic data warehouse design using a hybrid methodology," *Computer Science and Information Systems*, vol. 12, no. 1, pp. 135–160, 2015.
- [31] M. Mircea, A. I. Andreescu *et al.*, "Agile development for service oriented business intelligence solutions," *Database Systems Journal*, vol. 2, no. 1, pp. 43–56, 2011.
- [32] R. McHugh, F. Bilodeau, S. Rivest, Y. Bédard, and M. Michaud, "Analyse du potentiel d'une application solap pour une gestion efficace de l'érosion des berges en gaspésie îles-de-la-madeleine," *Proceedings of Géomatique*, 2006.

- [33] F. Ravat, O. Teste, and R. Tournier, "Analyse multidimensionnelle de documents via des dimensions olap," *Document numérique*, vol. 10, no. 2, pp. 85–104, 2007.
- [34] B. Sherrill, W. Eberle, and D. Talbert, "Analysis of student data for retention using data mining techniques," in *7th Annual National Symposium on Student Retention*, 2011, pp. 65–66.
- [35] C. Phipps and K. C. Davis, "Automating data warehouse conceptual schema design and evaluation." in *DMDW*, vol. 2. Citeseer, 2002, pp. 23–32.
- [36] S.-T. Li, L.-Y. Shue, and S.-F. Lee, "Business intelligence approach to supporting strategy-making of isp service management," *Expert Systems with Applications*, vol. 35, no. 3, pp. 739–754, 2008.
- [37] A. Girsang, S. Isa, A. Haris, K. Mandagie, L. Ariana, V. Ardinda *et al.*, "Business intelligence for product defect analysis," in *IOP Conference Series : Materials Science and Engineering*, vol. 598, no. 1. IOP Publishing, 2019, p. 012117.
- [38] A. Jadidi, M. A. Mostafavi, Y. Bédard, B. Long, and E. Grenier, "Using geospatial business intelligence paradigm to design a multidimensional conceptual model for efficient coastal erosion risk assessment," *Journal of coastal conservation*, vol. 17, no. 3, pp. 527–543, 2013.
- [39] D. Coelho, J. Miranda, F. Portela, J. Machado, M. F. Santos, and A. Abelha, "Towards of a business intelligence platform to portuguese misericórdias," *Procedia Computer Science*, vol. 100, pp. 762–767, 2016.
- [40] D. Ibragimov, K. Hose, T. B. Pedersen, and E. Zimányi, "Towards exploratory olap over linked open data—a case study," in *Enabling Real-Time Business Intelligence*. Springer, 2014, pp. 114–132.
- [41] N. Tryfona, F. Busborg, and J. G. Borch Christiansen, "starer : A conceptual model for data warehouse design," in *Proceedings of the 2nd ACM international workshop on Data warehousing and OLAP*, 1999, pp. 3–8.

- [42] A. Battaglia, M. Golfarelli, and S. Rizzi, "Qbx : a case tool for data mart design," in *International Conference on Conceptual Modeling*. Springer, 2011, pp. 358–363.
- [43] A. Delgado and A. Marotta, "Configurable web warehouses construction through bpm systems," *CLEI Electronic Journal*, vol. 19, no. 2, pp. 9–9, 2016.
- [44] F. S. Tseng and A. Y. Chou, "The concept of document warehousing for multi-dimensional modeling of textual-based business intelligence," *Decision Support Systems*, vol. 42, no. 2, pp. 727–744, 2006.
- [45] M. Serrano, J. Trujillo, C. Calero, and M. Piattini, "Metrics for data warehouse conceptual models understandability," *Information and Software Technology*, vol. 49, no. 8, pp. 851–870, 2007.
- [46] R. Wrembel and B. Bebel, "Metadata management in a multiversion data warehouse," in *Journal on data semantics VIII*. Springer, 2007, pp. 118–157.
- [47] D. Pushpal, "Implementation of time dimension in star schema design : A case study of birth registration data," *International Journal of Clothing Science and Technology*, vol. 6, no. 4, pp. 191–193, 2015.
- [48] D. Dori, R. Feldman, and A. Sturm, "From conceptual models to schemata : An object-process-based data warehouse construction method," *Information Systems*, vol. 33, no. 6, pp. 567–593, 2008.
- [49] B. Parmanto, M. Scotch, and S. Ahmad, "A framework for designing a healthcare outcome data warehouse," *Perspectives in Health Information Management/AHIMA, American Health Information Management Association*, vol. 2, 2005.
- [50] D. Sheta, E. Osama, and A. N. Eldeen, "Evaluating a healthcare data warehouse for cancer diseases," *arXiv preprint arXiv :1307.3448*, 2013.
- [51] I. Inuwa and N. Oye, "Designing of a data warehouse model for a university decision support system," *Information and Knowledge management*, vol. 5, no. 12, pp. 97–104, 2015.

- [52] R. Wrembel, “Data warehouse performance : selected techniques and data structures,” in *European Business Intelligence Summer School*. Springer, 2011, pp. 27–62.
- [53] I. Moalla, A. Nabli, L. Bouzguenda, and M. Hammami, “Data warehouse design from social media for opinion analysis : The case of facebook and twitter,” in *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*. IEEE, 2016, pp. 1–8.
- [54] L. Van Dyk and P. Conradie, “Creating business intelligence from course management systems,” *Campus-Wide Information Systems*, 2007.
- [55] F. Di Tria, E. Lefons, and F. Tangorra, “Hybrid methodology for data warehouse conceptual design by uml schemas,” *Information and Software Technology*, vol. 54, no. 4, pp. 360–379, 2012.
- [56] A. Chaves Carniel and T. L. Lopes Siqueira, “Querying data warehouses efficiently using the bitmap join index olap tool,” *CLEI Electronic Journal*, vol. 15, no. 2, 2012.
- [57] A. Vaisman and E. Zimányi, “Data warehouses : Next challenges,” in *European Business Intelligence Summer School*. Springer, 2011, pp. 1–26.
- [58] E. Malinowski and E. Zimányi, “Logical representation of a conceptual model for spatial data warehouses,” *GeoInformatica*, vol. 11, no. 4, pp. 431–457, 2007.
- [59] A. Vaisman and E. Zimányi, “Conceptual modeling for data warehouse and olap applications,” in *Encyclopedia of Data Warehousing and Mining, Second Edition*. IGI Global, 2009, pp. 293–300.
- [60] P. Giorgini, S. Rizzi, and M. Garzetti, “Goal-oriented requirement analysis for data warehouse design,” in *Proceedings of the 8th ACM international workshop on Data warehousing and OLAP*, 2005, pp. 47–56.
- [61] X. Hu and N. Cercone, “A data warehouse/online analytic processing framework for web usage mining and business intelligence reporting,” *International Journal of Intelligent Systems*, vol. 19, no. 7, pp. 585–606, 2004.

- [62] H. Brahmi, I. Brahmi, and S. B. Yahia, "Omc-ids : At the cross-roads of olap mining and intrusion detection," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2012, pp. 13–24.
- [63] S. H. A. El-Sappagh, A. M. A. Hendawi, and A. H. El Bastawissy, "A proposed model for data warehouse etl processes," *Journal of King Saud University-Computer and Information Sciences*, vol. 23, no. 2, pp. 91–104, 2011.
- [64] N. Prat, J. Akoka, and I. Comyn-Wattiau, "A uml-based data warehouse design method," *Decision support systems*, vol. 42, no. 3, pp. 1449–1473, 2006.
- [65] Peter Myers, Kate Follis, Maggie Sparkman, David Coulter, Kent Sharkey, Michael Blythe and Adam Saxton. (2019) Understand star schema and the importance for power bi. <https://docs.microsoft.com/en-us/power-bi/guidance/star-schema>.
- [66] Robert Sheldon. (2008) Sql server data warehouse cribsheet. <https://www.red-gate.com/simple-talk/sql/learn-sql-server/sql-server-data-warehouse-cribsheet/>.
- [67] Daryl Ung. (2019) Data warehousing : Basics of relational vs star schema data modeling. <https://medium.com/@daryl.ung/data-warehousing-basics-of-relational-vs-star-schema-data-modeling-75a68eeaf0e3>.
- [68] Nicole Janeway Bills. (2020) Comprehensive guide to the data warehouse. <https://towardsdatascience.com/data-warehouse-68ec63eef78>.
- [69] Iqbal Ahmed Alvi. (2018) Modeling your dimensional data warehouse : Star schema vs. snowflake schema. <https://datawarehouseinfo.com/data-warehouse-star-schema-vs-snowflake-schema/>.
- [70] Yazid Grim and Fleur-Anne Blain. (2005) Conception d'un entrepôt de données (data warehouse). <https://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse/>.
- [71] Sylvain Bourdette. (2001) Quels sont les buts recherchés et les enjeux d'un data-web ? <https://www.bourdette.com/Memoire/>.

- [72] Colleen M. Morrow. (2014) Exam prep 70-463 : Tables and schemas. <https://www.sqlservercentral.com/blogs/exam-prep-70-463-tables-and-schemas>.
- [73] Reza Rad. (2016) Data preparation; first and foremost important task in power bi. <https://radacad.com/data-preparation-first-and-foremost-important-task-in-power-bi>.
- [74] Enrico Besenyei. (2019) Star and snowflake schema in data warehousing. <https://www.eandbsoftware.org/star-and-snowflake-schema-in-data-warehousing/>.